Published for SISSA by Springer

Received: July 9, 2025
Revised: February 23, 2026
Accepted: May 11, 2026
Published: June 26, 2026

Experimental determination of BSM triple Higgs couplings at the HL-LHC with neural networks

Markus Frank,[1] **Sven Heinemeyer**,[a] **Margarete Mühlleitner**[b] **and Kateryna Radchenko**[a,c]

[a] *Instituto de Física Teórica (UAM/CSIC), Universidad Autónoma de Madrid, Cantoblanco, 28049, Madrid, Spain*

[b] *Institute for Theoretical Physics, Karlsruhe Institute of Technology, 76128 Karlsruhe, Germany*

[c] *Deutsches Elektronen-Synchrotron DESY, Notkestr. 85, 22607 Hamburg, Germany*

E-mail: markus.frank@x-ent.tech, Sven.Heinemeyer@cern.ch, margarete.muehlleitner@kit.edu, kateryna.radchenko@desy.de

Abstract: The shape of the Higgs potential is modified by the presence of additional scalar fields, as predicted in many Beyond-Standard-Model (BSM) scenarios. In such cases, deviations in the Higgs self-interactions — in particular the trilinear Higgs couplings — could serve to disentangle the physics beyond the Standard Model (SM). While the SM predicts only one trilinear Higgs coupling, extended scalar sectors allow for additional self-interactions that can manifest themselves in Higgs pair production, via the s-channel contribution of a heavy $\mathcal{CP}$-even scalar H. We present the first sensitivity study to such a BSM trilinear scalar coupling using machine learning. Specifically, we train a neural network on the invariant mass distributions of Higgs pair production at the HL-LHC to extract $\xi_H^t \times \lambda_{hhH}$, i.e. the product of the resonant H top-Yukawa coupling and the trilinear coupling of H to the two SM-like Higgses in the final state, hh. Assuming a hypothetical H mass of 450 GeV, we show that, depending on future experimental efficiencies and uncertainties, a determination of $\xi_H^t \times \lambda_{hhH}$ at the 10–20% level may be achievable by the end of the HL-LHC. We present a simple and more efficient alternative to classical statistical methods, proving the efficiency of neural networks for both hypothesis testing and parameter estimation, which outperforms conventional maximum likelihood methods in this context.

Keywords: Anomalous Higgs Couplings, Multi-Higgs Models, New Light Particles

ArXiv ePrint: 2506.18981

[1] Independent Researcher, Karlsruhe, Germany.

https://doi.org/10.1007/JHEP06(2026)274

Contents

1 Introduction

The discovery of a Higgs boson, h, at ATLAS and CMS in 2012, nearly 50 years after its prediction, was a milestone in high-energy physics [1, 2]. Within theoretical and experimental uncertainties this new particle is consistent with the existence of a Standard-Model (SM) Higgs boson at a mass of $\sim 125\,\text{GeV}$ [3]. Now the reconstruction of the Higgs potential is one of the main goals in particle physics in the years ahead. A primary objective is to establish the trilinear Higgs coupling (THC), λ_{hhh}, with a high level of precision. This task can partially be accomplished in the High Luminosity phase of the Large Hadron Collider (HL-LHC) [4]. While no sign of Beyond-Standard-Model (BSM) physics was yet discovered at the LHC, the measurements of Higgs-boson production and decay rates, which are known experimentally to a precision of roughly $\sim 10{-}20\%$ [5, 6], leave ample room for BSM interpretations [7]. Consequently, the possibility of additional scalar states present in nature that have escaped detection so far remains an enticing research avenue. If a new signal is established in the

JHEP06(2026)274

current or upcoming runs of the LHC, the next question will be to determine the couplings of the new scalar states at the highest possible level of precision. In particular, the THCs of these new scalar states are crucial for a full reconstruction of the Higgs potential.

One important reason to consider a non minimal scalar sector is the possibility to accommodate a strong first order electroweak phase transition (SFOEWPT), a necessary condition for electroweak baryogenesis [8–11]. It has been found that the SM does not feature such a transition with the mass of the discovered Higgs being $\sim 125\,\text{GeV}$ [12]. However, models with extended scalar sectors can accommodate a SFOEWPT, cf. e.g. for models with a second Higgs doublet [8, 9, 11, 13–19], which in turn may lead to a detectable gravitational waves signal [20–22].

A very well motivated extension of the SM that could yield such a signal is the Two-Higgs-Doublet Model (2HDM) [23–26], that extends the field content of the SM by an additional complex Higgs doublet. This results after electroweak symmetry breaking in five physical scalar fields, two $\mathcal{CP}$-even scalar fields, h and H, where by convention $m_h < m_H$ (we assume $m_h \sim 125\,\text{GeV}$), one $\mathcal{CP}$-odd field, A, and one charged Higgs pair, $H^\pm$. If such a scenario was realized in nature, not only the measurement of the SM-like THC, λ_{hhh}, would be needed, but also experimental sensitivity to the remaining THCs among all the scalars in the theory will be necessary to determine the shape of the Higgs potential.

Triple Higgs couplings enter at leading order in the main process of di-Higgs production at the LHC, given by gluon fusion. Consequently, this process is of particular interest in the experimental determination of the THCs. Within the SM, two Feynman diagrams contribute to hh production: via a top-quark triangle a light Higgs, h, is produced in the s-channel with the subsequent decay into hh, and the diagram where the two Higgs bosons are radiated from a top-quark box. In the 2HDM there is one additional contribution to this process: via a top-quark triangle a heavy Higgs, H, is resonantly produced. Consequently, hh production would provide access not only to λ_{hhh}, but also to λ_{hhH}. The relevant quantity for the H contribution is $\xi_H^t \times \lambda_{hhH}$, where ξ_H^t parametrizes the top-Yukawa coupling of the heavy Higgs boson.

In view of an ever increasing availability of experimental data, "old school" data treatment such as χ^2 tests become insufficient when dealing with the determination of BSM THCs at HL-LHC. The broad selection of newly available tools for data analysis has been used extensively by the high-energy physics community in the past years [27, 28]. In particular, in experimental particle physics, modern machine learning techniques have proven to be extremely useful in, e.g., the event selection. In particle physics phenomenology a wide range of studies has also started to employ neural networks and machine learning techniques as a tool to learn about the underlying physics of nature.

In the context of di-Higgs production, machine learning techniques have recently been applied in order to enhance the sensitivity to Higgs couplings in non-resonant hh production. For example, ML architectures trained on multi-dimensional kinematic information have been shown to improve constraints on the trilinear Higgs self-coupling under realistic LHC conditions [29]. Interpretable ML approaches have also been employed to extract information on Higgs self-interactions and light-quark Yukawa couplings using di-Higgs observables and have been shown to outperform traditional cut based analyses [30].

JHEP06(2026)274

In this work, we target resonant di-Higgs production in the context of the 2HDM, with the primary goal of extracting the product of the heavy scalar top Yukawa coupling and the trilinear Hhh scalar coupling, $\xi_H^t \times \lambda_{hhH}$, from the invariant mass distribution of the di-Higgs pair, m_{hh}, in the process $gg \to hh$. To this end, we will analyze the performance of machine learning techniques, more specifically a neural network (NN), in the determination of such couplings using anticipated HL-LHC data for di-Higgs production, as opposed to the conventional statistical data analysis methods that are insufficient for the extraction of BSM THCs [31]. To the best of our knowledge, this work constitutes the first analysis attempting to quantitatively extract BSM THCs from (HL-)LHC data.

This paper is organized as follows. First, in section 2, we give a very brief introduction of the 2HDM where we settle our notation and specify the analyzed parameter region as well as the constraints applied in the model. In section 3 we provide a classical statistical analysis of the data to determine $\xi_H^t \times \lambda_{hhH}$. In section 4 we give details of the methodology followed in our NN analysis, in particular in the chosen data sets and type of network. We show our results in section 5, where a comparison between both methods and future prospects for the determination of $\xi_H^t \times \lambda_{hhH}$ are analyzed. We state our conclusions in section 6.

2 The 2HDM

Our analysis is performed in the $\mathcal{CP}$-conserving 2HDM with a softly broken $\mathbb{Z}_2$ symmetry [23–26]. This discrete symmetry is imposed in order to avoid flavor changing neutral currents at tree level, which are experimentally found to be very suppressed. The 2HDM Higgs sector is comprised of two complex Higgs doublets, leading after EWSB to five physical Higgs bosons, two $\mathcal{CP}$-even states h and H, one $\mathcal{CP}$-odd state A and two charged states $H^\pm$. The mass hierarchy of the $\mathcal{CP}$-even particles is by convention $m_H > m_h$, and we choose to identify the light state h with the experimentally measured state at $\approx 125\,\text{GeV}$. The electroweak precision observables constrain the masses of the remaining scalars such that either $m_{H^\pm} \approx m_H$ or $m_A \approx m_H$. Therefore, for simplicity, in the following analysis we will assume $m_\phi \equiv m_H = m_A = m_{H^\pm} > m_h$. (As will become clear later, the masses of the $\mathcal{CP}$-odd and charged Higgs bosons play a negligible role in our analysis.) The imposed $\mathbb{Z}_2$ symmetry is allowed to be softly broken by a dimensionful mass term m_{12}^2, which is a free parameter of the model. Furthermore, there are two mixing angles α and β that diagonalize the $\mathcal{CP}$-even and $\mathcal{CP}$-odd/charged sectors, respectively. It is convenient to use the parameterization $\tan\beta \equiv t_\beta$, which is the ratio of the vacuum expectation values of the two doublets, i.e. $t_\beta = v_2/v_1$, and $c_{\beta-\alpha} \equiv \cos(\beta-\alpha)$, which is zero in the so-called alignment limit [32]. In this limit all the couplings of the scalar h to fermions and gauge bosons have the same values as predicted in the SM and in particular $\lambda_{hhH} = 0$.

In this set-up, the free parameters describing the tree-level Higgs sector of the model, given that $m_h \approx 125\,\text{GeV}$ and $v = \sqrt{v_1^2 + v_2^2} \approx 246\,\text{GeV}$, can be chosen as

$$m_\phi,\ m_{12}^2,\ t_\beta,\ c_{\beta-\alpha}\ . \tag{2.1}$$

The imposition of a discrete $\mathbb{Z}_2$ symmetry leads four different 2HDM types, depending on the extension of the $\mathbb{Z}_2$ symmetry to the fermion sector. This determines the Yukawa

JHEP06(2026)274

couplings of the Higgs bosons to the fermions. In this work we will focus only on scenarios in the so called "Type I" 2HDM, although the type of the 2HDM plays a subleading role and we can safely generalize our results, e.g. the parameter in which we are interested, $\xi_H^t \times \lambda_{hhH}$, is the same in all four Yukawa types. Relevant to our analysis here will be the Yukawa couplings of the $\mathcal{CP}$-even Higgs bosons, which enter in the 2HDM Lagrangian as

$$\mathcal{L} = -\sum_{f=u,d,l} \frac{m_f}{v} \left[\xi_h^f \bar{f} f h + \xi_H^f \bar{f} f H \right], \tag{2.2}$$

where m_f are the fermion masses and $\xi_{h,H}^f$ are the fermionic Yukawa coupling modifiers, which in Type I are equal for all three generations of up-type quarks, down-type quarks and leptons. Since the Higgs couples most strongly to the top quark, we will consider only the top Yukawas in the following ($f = t$). In all four types the Yukawa coupling modifiers for the top quark are given by

$$\xi_h^t = s_{\beta-\alpha} + c_{\beta-\alpha} \cot\beta, \quad \xi_H^t = c_{\beta-\alpha} - s_{\beta-\alpha} \cot\beta. \tag{2.3}$$

And the tree level expressions of the THCs (which are independent of the Yukawa type) are given by

$$\lambda_{hhh} = \frac{1}{2v^2} \left\{ m_h^2 s_{\beta-\alpha}^3 + \left(3m_h^2 - 2M^2\right) c_{\beta-\alpha}^2 s_{\beta-\alpha} + 2\cot 2\beta \left(m_h^2 - M^2\right) c_{\beta-\alpha}^3 \right\}, \tag{2.4}$$

$$\begin{aligned} \lambda_{hhH} = -\frac{c_{\beta-\alpha}}{2v^2} \Big\{ & \left(2m_h^2 + m_H^2 - 4M^2\right) s_{\beta-\alpha}^2 + 2\cot 2\beta \left(2m_h^2 + m_H^2 - 3M^2\right) s_{\beta-\alpha} c_{\beta-\alpha} \\ & - \left(2m_h^2 + m_H^2 - 2M^2\right) c_{\beta-\alpha}^2 \Big\}, \end{aligned}$$

where $s_{\beta-\alpha} \equiv \sin(\beta-\alpha)$ etc. and M^2 is defined as

$$M^2 = \frac{m_{12}^2}{s_\beta c_\beta}. \tag{2.5}$$

It should be noted here that we do not consider loop corrections to the THC, although these can be relevant in di-Higgs production [33–37]. We focus on the question whether, given the predictions for the invariant mass distribution in Higgs pair production through gluon fusion within a theoretical framework, the experimental determination of the couplings is possible. In the hypothetical case of a future detection of a heavy Higgs resonance in di-Higgs production, such a framework must be improved with higher order corrections. However, this is not expected to alter the performance of the NN analysis.

2.1 Higgs pair production

In the framework of the 2HDM, Higgs pair production has been widely studied in the literature for many years (cf. e.g. refs. [38–40] for the Minimal Supersymmetric Standard Model (MSSM) that shares the 2HDM Higgs sector with Yukawa Type II, and e.g. refs. [7, 17, 31, 37, 41–45] for the 2HDM). The leading-order diagrams for gluon fusion into a SM-like Higgs pair in the 2HDM are shown in figure 1. The box (left) and the h exchange in the s-channel (middle) comprise the continuum contribution that is present in the SM (modulo changes in the

JHEP06(2026)274

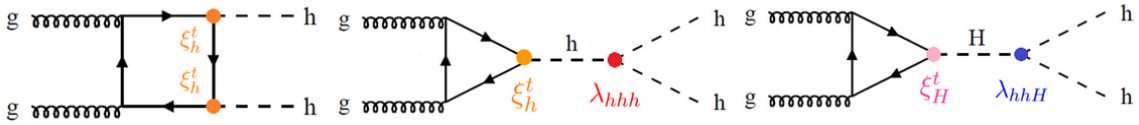

Figure 1. Leading-order diagrams to 2HDM SM-like Higgs pair production in the gluon fusion process at hadron colliders.

Higgs-boson couplings w.r.t. their SM values). The resonant diagram (right) involves the heavy $\mathcal{CP}$-even Higgs-boson, H, exchanged in the s-channel. The cross section given by the continuum diagrams is proportional to the SM-like top Yukawa coupling ξ_h^t and the trilinear Higgs self-interaction λ_{hhh}. The resonant diagram, entering in the full calculation of the di-Higgs production cross section in the 2HDM is proportional to the product of the heavy Higgs-boson top-Yukawa coupling ξ_H^t and the THC λ_{hhH}. The interference of this diagram with the continuum contribution creates a dip-peak or peak-dip structure in the m_{hh} distribution, located at $\sim m_H$, where the type of the structure depends on the sign of the product of $(\xi_H^t \times \lambda_{hhH})$ [31]. It was furthermore shown in ref. [31] that in parameter regions close to the alignment limit, i.e. with the properties of the light $\mathcal{CP}$-even Higgs boson in good agreement with the LHC rate measurements, and with $m_H \lesssim 1000\,\text{GeV}$ the resonant H contribution can modify the m_{hh} distribution substantially. This suggests that an extraction of $\xi_H^t \times \lambda_{hhH}$ via the measurement of the invariant mass distribution m_{hh} of the di-Higgs production cross section may be possible.

For the theoretical prediction of the Higgs pair production cross section we use the code `HPAIR` [38, 39] adapted for the 2HDM, which also computes the invariant mass distribution of the two light Higgses hh in the final state, m_{hh}. It allows to include next-to-leading-order QCD corrections in the heavy-top limit.

2.2 Model constraints

In order to analyze viable 2HDM parameter points, the relevant constraints on the model have to be taken into account. In this work we have applied the most relevant constraints coming from theoretical considerations and previous experimental measurements. In order to take these constraints into account we have used the public tool `thdmTools` [46]. In particular, we have applied bounds for perturbative unitarity (we choose an upper bound for the eigenvalues of the scalar four-point scattering matrix of 8π), vacuum stability, constraints from electroweak precision observables, flavor constraints and constraints from direct BSM Higgs-boson searches at the LHC, Tevatron and LEP, as well as the compatibility with the signal strength of the Higgs-boson at $\sim 125\,\text{GeV}$. The latter two are taken into account through the interface to `HiggsTools` [47] (which merged and updated `HiggsBounds` [48–52] and `HiggsSignals` [53–55]).

2.3 Benchmark scenario

In our study we will concentrate on one exemplary benchmark plane, which is defined in the 2HDM Type I. In our analysis we will only include parameter points that are in agreement with the constraints discussed in the previous subsection. The example is chosen to be

JHEP06(2026)274

representative of a scenario with resonant di-Higgs production: we have set $m_\phi = 450\,\text{GeV}$, which is assumed to be measured in a different process, i.e. with a different, less complex final state, where later the corresponding experimental uncertainty will be considered. We have taken t_β and $c_{\beta-\alpha}$ as free parameters, assuming that their determination in the future will be more complicated. m_{12}^2 is either fixed via[1]

$$m_{12}^2 \equiv m_H^2 c_\alpha^2 / t_\beta \,, \tag{2.6}$$

or taken as a free parameter (see below). For this parameter the experimental prospects at the HL-LHC are unclear [56], and a variation of m_{12}^2 as a third free parameter is the most conservative approach. In figure 2 we show our benchmark plane for m_{12}^2 fixed according to eq. (2.6). The colored region is allowed by all the applied constraints while the white region is excluded by either of them. In particular, the lower bound on t_β is set by flavor constraints coming from $B \to X_s\gamma$ [57] searches, the bound on negative $c_{\beta-\alpha}$ comes from requiring the stability of the vacuum at tree level (using the conditions defined in ref. [58]) and for positive $c_{\beta-\alpha}$ and large t_β from requiring perturbative unitarity at tree level (using the conditions defined in ref. [59]). Additionally, for lower t_β values and larger $c_{\beta-\alpha}$, searches for resonant di-Higgs production, in particular, the combination of searches in ref. [60], provide the most stringent constraints. For intermediate values of $c_{\beta-\alpha}$ the most stringent bound comes from the requirement of compatibility of the Higgs boson at 125 GeV signal strengths with the experimental measurements (we include in particular the combination of all the measurements included in the develop branch of `HiggsSignals` [47]). The color coding indicates the product of the couplings $\xi_H^t \times \lambda_{hhH}$ that we will later aim to extract. The black lines denote, where we find $\xi_H^t \times \lambda_{hhH} = 0$, and the three crosses mark benchmark points discussed below. It can be observed that $\xi_H^t \times \lambda_{hhH} = 0$ for $c_{\beta-\alpha} = 0$, i.e. in the alignment limit, as expected. However, the value of $\xi_H^t \times \lambda_{hhH} = 0$ is also found along a curve for $\lambda_{hhH} = 0$ at $c_{\beta-\alpha} = 0.02 - 0.12$ and $t_\beta = 8 - 50$, as well as for $\xi_H^t = 0$ at $c_{\beta-\alpha} = 0.03 - 0.10$ and $t_\beta = 15 - 50$. The values for $\xi_H^t \times \lambda_{hhH}$ (in various set-ups) together with the corresponding m_{hh} distributions are going to be the input for a neural network as will be described below.

2.4 Invariant mass distributions

The core of our analysis will be the connection between the value of $\xi_H^t \times \lambda_{hhH}$ and the shapes of the corresponding m_{hh} distributions. The shape of the invariant mass distribution of di-Higgs production has been analyzed with great detail in the past. In ref. [61] even the possibility of classifying different kinds of distributions through a neural network approach was investigated. The projected shapes were classified in the region of the coupling parameters space in an EFT approach, allowing for the identification of deviations in the couplings involved in the SM gluon fusion di-Higgs production trough the invariant mass shape analysis. However, since no further BSM state was assumed, there was no investigation of the role of a BSM THC.

In previous works we have demonstrated that di-Higgs invariant mass distributions at the HL-LHC could possibly provide access to the BSM THC λ_{hhH} [31].[2] In particular, assuming that a resonant scenario is realized, i.e. the contribution of H is sizable, the sign of the

[1] This relation ensures that perturbative unitarity is fulfilled for large t_β.

[2] Corresponding analyses for future e^+e^- colliders can be found in refs. [62–64].

JHEP06(2026)274

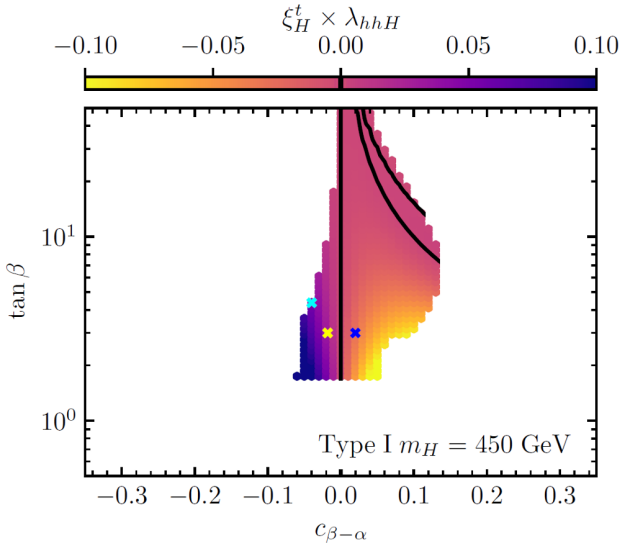

Figure 2. $\xi_H^t \times \lambda_{hhH}$ in the example benchmark plane in the Type I 2HDM with $m_\phi = 450\,\text{GeV}$ and m_{12}^2 fixed via eq. (2.6). Black lines are located at $\xi_H^t \times \lambda_{hhH} = 0$, either because $\lambda_{hhH} = 0$ (in the alignment limit, i.e. the vertical line, and the lower curved line on the right upper corner) or because $\xi_H^t = 0$ (the upper line on the right corner). The blue cross indicates an example point whose m_{hh} distribution is displayed in figure 3. The yellow cross features an example with roughly the opposite sign value of $\xi_H^t \times \lambda_{hhH}$ and its distribution is displayed in figure 4. The cyan cross is an example point for which different predictions of the NN will be analyzed further in figure 14.

product $\xi_H^t \times \lambda_{hhH}$ would determine the resonant structure at $m_{hh} \approx m_H$. More precisely, the structure would be dip-peak for an overall negative sign and peak-dip for an overall positive sign, assuming that no further BSM effects arise from e.g. loop corrections to the λ_{hhh} and λ_{hhH} trilinear Higgs coupling (see refs. [31, 37] for further details). However, these effects are partially washed out by experimental uncertainties, as we briefly review in the next subsection.

2.5 Experimental challenges

An example for the theory prediction for the m_{hh} distribution, which is the key quantity for our analysis, is shown in the left plot of figure 3. The parameter point is taken from figure 2, marked there with a blue cross, with $c_{\beta-\alpha} = 0.02$ and $t_\beta = 3$. For this point, we find a value of $\xi_H^t \times \lambda_{hhH} = -0.029$. The solid blue line shows the full m_{hh} distribution, while the dashed black line is calculated from the two continuum diagrams only, i.e. leaving out the H-resonance contribution. The dip-peak structure at $m_{hh} = m_H = 450\,\text{GeV}$ is clearly visible.

However, the theory prediction is subject to experimental uncertainties, see ref. [31] for a detailed analysis, used to obtain our anticipated uncertainties. The first assumption is that the analysis can be done with a (QCD) background subtracted data sample. While this may appear optimistic now, we expect that at the end of the HL-LHC run this will be standard practice in the di-Higgs measurements of ATLAS and CMS (see, e.g., ref. [65] for a similar treatment in the vector boson fusion $h \to b\bar{b}$ production channel after the Higgs discovery). Correspondingly, we do not assign an additional uncertainty on the background subtraction. We note, however, that residual systematics in the background shape could, in

JHEP06(2026)274

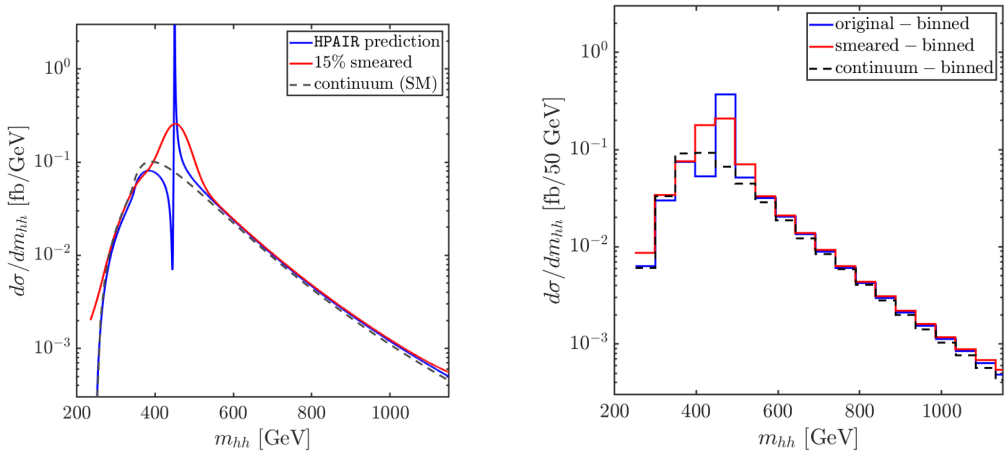

Figure 3. Smearing (left) and binning (right) applied to the invariant mass distribution prediction for an example benchmark point in the 2HDM, marked in blue in figure 2. The dashed lines indicate the continuum contribution.

principle, mimic the subtle dip-peak structure of our signal and may therefore have a larger impact than in the $h \to b\bar{b}$ case. Our approach is intended as a proof-of-concept, and this assumption should be kept in mind when interpreting the results. The first experimental uncertainty that we do take into account is the uncertainty in the m_{hh} measurement. Each point of the m_{hh} distribution is "smeared" with a Gauss distribution with a width of 15%, as shown in the red curves in figure 3. It can be observed in the solid red line in the left plot that the smearing significantly dilutes the dip-peak structure and only a very small broader dip, followed by a broad "hump", remains.

On top of that, experimental data is gathered in bins, owing to the finite energy resolution, which may hide some of the effects that would be visible if the resolution was perfect. The bin size is set to 50 GeV, resulting in 16 bins (which serve later as input for our NN analysis). In the right plot of figure 3 the solid blue curve shows the binned distribution for the unsmeared blue curve of the left plot. The solid red curve is the binned distribution of the smeared red curve in the left plot. Both are compared to the smeared and binned continuum distribution, i.e. the distribution neglecting the resonance, shown as black dashed curve. It can be observed that in this case the dip-peak structure around the resonance persists only very mildly (when compared to the SM prediction) after both smearing and binning have been taken into account. If one further compares the red curve of figure 3 with a different one that yields a positive value of $\xi_H^t \times \lambda_{hhH}$, the peak-dip/dip-peak structure cannot be resolved optically anymore. This is demonstrated in figure 4, where we show such a comparison, taking a point from figure 2 with $t_\beta = 3$, $c_{\beta-\alpha} = -0.018$. This yields a value of $\xi_H^t \times \lambda_{hhH} = 0.0304$ (this point is marked with a yellow cross in figure 2), which is roughly the same magnitude but opposite sign w.r.t. the one giving rise to the red line. This poses a challenge on the experimental access to $\xi_H^t \times \lambda_{hhH}$ that we will try to address in the following sections.

It should be noted that the numbers chosen for smearing (15%) and binning (50 GeV) were recommended within the LHC Higgs working group and are expected to correspond to a reasonable (realistic) experimental set up.

JHEP06(2026)274

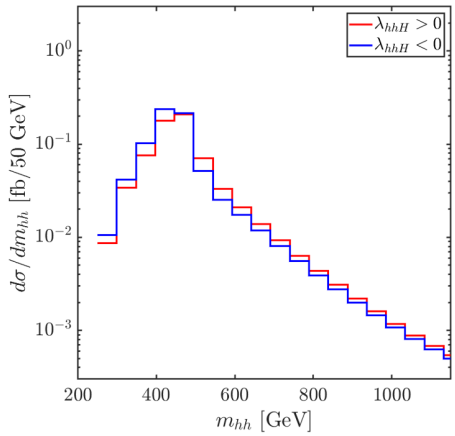

Figure 4. Comparison of two benchmark points yielding a positive (blue) and negative (red) value of $\xi_H^t \times \lambda_{hhH}$ with $\xi_H^t \times \lambda_{hhH} \approx \pm 0.03$. The blue (red) curve corresponds to the point marked by a blue (yellow) cross in figure 2.

2.6 Statistical uncertainties

The final location of the data points of an experimental distribution will also be affected by an unavoidable uncertainty of the data itself. This consists of three components, a statistical uncertainty, a theoretical uncertainty (from unknown higher-order corrections) and a systematic uncertainty. Here we will only consider the first one because the latter two are harder to estimate and in an optimistic scenario should be subdominant. In order to obtain the statistical uncertainty we have to determine the expected number of events in each bin of the m_{hh} distribution. Experimental efficiencies are only available for definite final states, i.e. taking into account the decay of the hh system. A combination of various final states, on the other hand, is an experimental analysis on its own and thus goes far beyond the scope of our NN focused work. The most promising channel in this regard (i.e. the one producing the largest number of events taking into account experimental efficiencies) is $gg \to hh \to b\bar{b}\,b\bar{b}$. We have calculated the expected number of events in this channel in each bin as:

$$N_i \;=\; \sigma_i(gg \to hh) \;\times\; \mathcal{L} \;\times\; \mathrm{BR}^2(h \to b\bar{b}) \;\times\; \epsilon, \tag{2.7}$$

where $\sigma_i(gg \to hh)$ is the differential production cross section in each m_{hh} bin i times the size of the bin (in our case 50 GeV), $\mathcal{L} = 3000\,\mathrm{fb}^{-1}$ is the integrated luminosity expected at the end of the HL-LHC, $\mathrm{BR}(h \to b\bar{b}) = 0.5841$ is the branching ratio of the decay of a SM Higgs boson into a pair of bottom quarks. Finally, $\epsilon \equiv \epsilon_{\mathrm{TOT}}\,\epsilon_{\mathrm{SR}}$ is the detector efficiency. Here ϵ_{TOT} is the preselection efficiency, in this case it is the number of events with ≥ 2 b-tagged jets over the total number of events ($N = \sigma \times \mathrm{BR} \times \mathcal{L}$), and ϵ_{SR} is the efficiency of the signal region (SR), i.e. the number of di-Higgs events out of the preselected events. In practice, N_i, the number of events in each bin, is expected to follow a Poisson distribution with the mean given by eq. (2.7). For the $b\bar{b}\,b\bar{b}$ channel we took the efficiencies from figure 3 (left plot for $s = 0$) in ref. [66] for $m_H = 450\,\mathrm{GeV}$. They are 17.3% and 1% for the total

JHEP06(2026)274

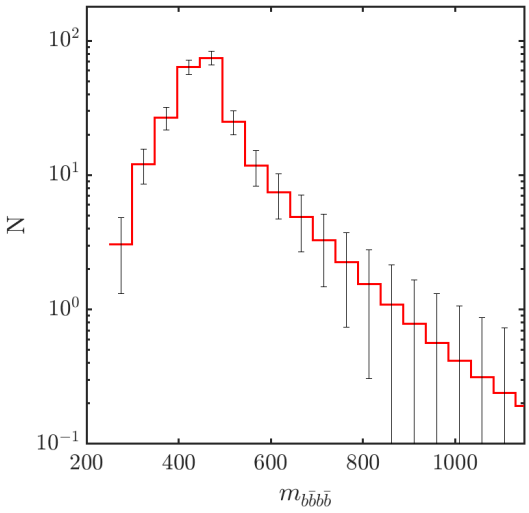

Figure 5. Example number of events in the $b\bar{b}\,b\bar{b}$ channel with statistical error bars (see text) for the point marked by the blue cross in figure 2.

and the signal region efficiency, respectively. The statistical uncertainty in bin i is then given by $\sqrt{N_i}$, where N_i is given by eq. (2.7).[3]

In figure 5 we show one example of the number of events in the $b\bar{b}\,b\bar{b}$ channel for the parameter point of the benchmark scenario with $m_H = 450\,\text{GeV}$ marked by the blue cross in figure 2. The statistical uncertainties are shown as black error bars in each m_{hh} bin. These statistical uncertainties will be interpreted as one standard deviation of a Poisson distribution around the (theoretical) true value, which is the mean of the Poisson distribution.

3 Statistical analysis

In this section we investigate the determination of $\xi_H^t \times \lambda_{hhH}$ with classical methods, i.e. not using NN. Statistically, we can perform an hypothesis test of the SM. We compare it to the null hypothesis (H_0), which is $\xi_H^t \times \lambda_{hhH} = 0$ (as realized in the SM), to an alternative hypothesis (H_1), which is the full 2HDM prediction including the H-resonance contribution. The deviation is then evaluated for our parameter of interest, $\xi_H^t \times \lambda_{hhH}$. It should be kept in mind that we take the 2HDM just as a plausible model where such non-zero values of the BSM trilinear couplings can be accommodated. Our results should be considered more general, however, i.e. a similar analysis can be performed in any model containing a heavy $\mathcal{CP}$-even Higgs resonance contributing to di-Higgs production at the HL-LHC. For the aforementioned hypothesis test we compute the likelihood ratio test between the null model (which gives rise to no resonance and $\xi_H^t \times \lambda_{hhH} = 0$), and the alternative model (where $\xi_H^t \times \lambda_{hhH} \neq 0$).

[3]Using a more realistic requirement for the efficiencies corresponding to the b-tagging of 4 jets reduces the signal efficiency from 17.3% to 5.8%. We will comment in section 5.4 on how much our results are affected by this assumption. However, on the other hand, we note that we do not consider an additional factor of 2 coming from a possible combination of both ATLAS and CMS data (corresponding to an effective doubling of the luminosity).

JHEP06(2026)274

As the observed values in each bin are modeled as samples from a Poisson distribution, the likelihood of each hypothesis is given by

$$L(H_0) = \prod_{i=0}^{15} \frac{{m_i}^{n_i} e^{-m_i}}{n_i!} \quad \text{and} \quad L(H_1) = \prod_{i=0}^{15} \frac{{n_i}^{n_i} e^{-n_i}}{n_i!}, \tag{3.1}$$

where m_i is the predicted mean count of the number of events in each bin in the 2HDM, and n_i is the observed count from experimental data. The likelihood ratio (LR) test statistic, λ_{LR}, is defined as

$$\lambda_{\text{LR}} = -2 \ln \left(\frac{L(H_0)}{L(H_1)} \right) = 2 \sum_{i=0}^{N_{\text{bins}}} \left(n_i \ln \frac{n_i}{m_i} + m_i - n_i \right). \tag{3.2}$$

According to Wilk's theorem, λ_{LR} asymptotically follows a χ^2 distribution, provided that the sample size is large enough: as a rule of thumb each n_i and m_i should be larger than 5. This theorem allows to directly compute a p-value as

$$p = P(\chi^2 > \lambda_{\text{LR}}) \,. \tag{3.3}$$

The p-value is the probability (P) of obtaining a result at least as extreme as the data, assuming the null hypothesis is true. Typically values of $p < 0.05$ (roughly corresponding to 95% CL) suggest evidence against the null hypothesis. In order to correctly apply Wilk's theorem, we need to ensure that the number of events in each bin is "large enough". Therefore we aggregate some bins together in order to have better statistics. We show an example of this aggregation in figure 6, where two sample distributions from figure 2 are shown: one with a small product of couplings, $\xi_H^t \times \lambda_{hhH} = 0.0188$ (orange), and one with a large value, $\xi_H^t \times \lambda_{hhH} = 0.0553$ (green). They are compared to the SM distribution (blue), which is equivalent to the alignment limit. On the left we show the distributions with the original 50 GeV binning and on the right the aggregated bins that lead to a particle count of at least 4 events, so that the statistical tests can be performed.[4] In particular, the first and second bin are added up together, and correspond to the first bin on the right. The third, fourth, fifth and sixth bin are unchanged and correspond to the bins 2, 3, 4 and 5 on the right. All the other bins are summed in bin number 6 on the right. We use the aggregated bins for the statistical analysis performed below, for the NN we keep the extended 16-bin sample, as the statistical considerations discussed above do not play a role in the NN.

We verified the robustness of our results by repeating the classical inference using the full 16-bin Poisson likelihood with toy-MC calibration, and by repeating the NN analysis with aggregated bins. The resulting differences are minor, confirming that the comparison is not biased by binning choices. In practice, we generated $N_{\text{toys}} = 5 \cdot 10^4$ pseudo-experiments for each parameter point. The fluctuations of the data are drawn using a random Poisson distribution, and the likelihood ratio is computed according to eq. (3.2).

The p-values for the benchmark plane shown in figure 2 are depicted in figure 7. For the parameter points for which the computed p-values are less than 0.05 (colored in blue), the

[4] We have performed a statistical analysis only using the bins which initially have $n_i \geq 5$ events. The result for that case is similar as the one using aggregated bins.

JHEP06(2026)274

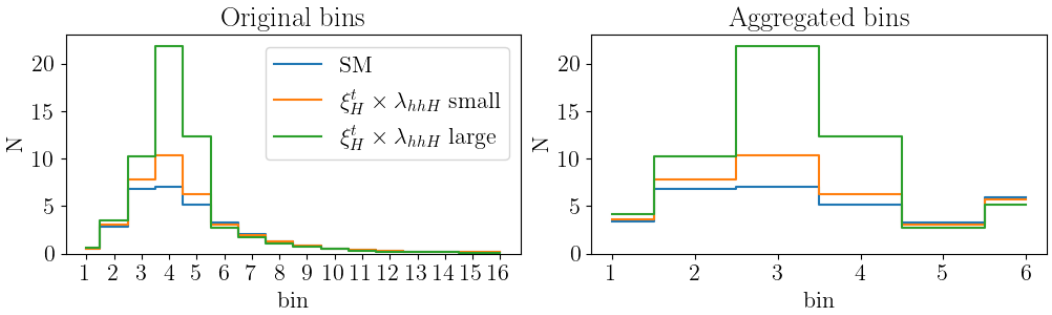

Figure 6. Sample distributions for the number of events with the original 50 GeV binning (left) and the aggregated bins (right). Shown are the distributions for the SM (blue), and two examples with a small value of $\xi_H^t \times \lambda_{hhH} = 0.0188$ (orange), and with a large value, $\xi_H^t \times \lambda_{hhH} = 0.0553$ (green).

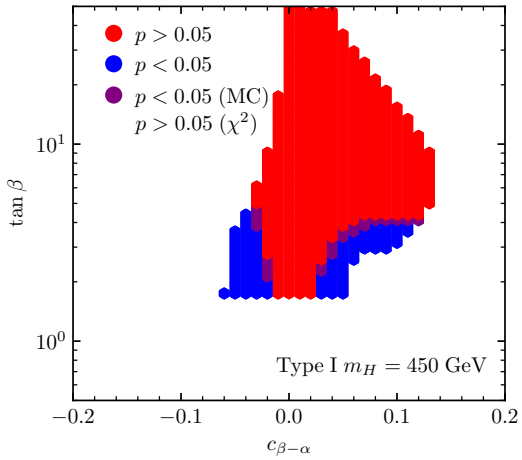

Figure 7. p-value in the example benchmark plane. The red regions shows the parameters points where the p-value is larger than 0.05, i.e. the null hypothesis cannot be rejected. The blue colored points represent the region where the null hypothesis can be tested with classical methods and a deviation from $\xi_H^t \times \lambda_{hhH} = 0$ is statistically significant. The purple region represents the parameter space where the toy-MC fails to reject the null hypothesis but the χ^2 test does not.

null hypothesis can be rejected. This method is not used for parameter estimation, but "just" for hypothesis testing, and the bottom line is that the distributions of the points within the red regions are compatible with the SM at 95% CL.

As a side note, the square root of the a χ^2-distributed variable, in our case λ_{LR}, can be interpreted in terms of standard deviations from the null hypothesis. In particular if $\sqrt{\lambda_{\text{LR}}} > 1.96$, the conclusion would be the same as if $p < 0.05$, meaning such a point is statistically far from the null hypothesis. However, this approach only works exactly for 1 degree of freedom, and gets worse for our case, where we have 6 degrees of freedom (i.e. the $N_{\text{bins}} = 6$ after aggregation). For this reason, we show in figure 7 the p-values for the benchmark plane defined in figure 2, rather than the significance.

When using the toy-MC, we find that the resulting p-values are generally slightly less aggressive than those obtained using the asymptotic χ^2 approximation with aggregated bins, i.e. fewer BSM parameter points are excluded at a fixed confidence level, as can be observed

JHEP06(2026)274

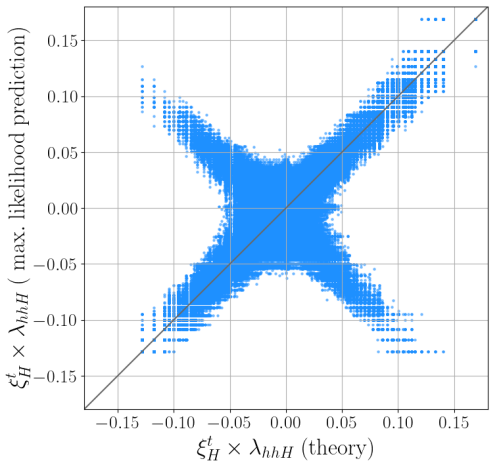

Figure 8. Predicted value of $\xi_H^t \times \lambda_{hhH}$ with the classical maximum likelihood estimation method versus the theoretical value of $\xi_H^t \times \lambda_{hhH}$.

from the purple region of figure 7. This behaviour is expected in the low-statistics regime, where the conditions underlying Wilks' theorem are not satisfied. When several bins contain small event counts, the asymptotic χ^2 approximation tends to underestimate the true width of the test-statistic distribution and effectively overcounts the available degrees of freedom, leading to p-values that are biased low. As a result, the asymptotic treatment typically overestimates the exclusion power compared to an exact Poisson likelihood with toy-MC calibration. This result further reinforces the need to use alternative approaches such as the NN analysis that we perform. In the following we will use the results of toy-MC p-value.

The p-value is a method used for hypothesis testing, and cannot really serve the purpose of our problem. We want to estimate a parameter $\xi_H^t \times \lambda_{hhH}$ from a distribution. Classically, for parameter estimation, the most precise method is the maximum likelihood estimation (MLE). The likelihood of each BSM scenario is defined as in eq. (3.1). We calculate $L(H_0)$ for each set of m_i (all points in figure 2) and find the one that maximizes L given the experimentally observed count n_i. This would be the most likely parameter point to produce the observed distribution. This method is only meaningful if the input data is assumed to be Poisson distributed around the true value. Otherwise the theoretically predicted point will trivially be the best fit. Since this procedure does not rely on Wilks' theorem, we used the full 16-binned sample distributions.

Figure 8 shows the predicted parameter $\xi_H^t \times \lambda_{hhH}$ with the MLE method versus the theoretical $\xi_H^t \times \lambda_{hhH}$. In total, our sample consists of 2048 m_{hh} distributions, smeared according to their Poisson distribution. We see that the method fails to correctly predict the sign of the $\xi_H^t \times \lambda_{hhH}$ couplings, as we expected from the invariant mass distributions. The 'X' shape in figure 8 explicitly demonstrates that while the method manages to determine if $\xi_H^t \times \lambda_{hhH} \neq 0$, the sign of $\xi_H^t \times \lambda_{hhH}$ cannot be predicted. These limitations necessitate the exploration of other techniques, such as neural networks, to achieve the required sensitivity and precision for the numerical estimation of $\xi_H^t \times \lambda_{hhH}$.[5]

[5]Classical methods can still be competitive in the determination of other BSM couplings such as the $\mathcal{CP}$-violating top Yukawa interaction [67].

JHEP06(2026)274

4 Determination of $\xi_H^t \times \lambda_{hhH}$ with NNs

In this work, we focus on the following problem: the extraction of the value of $\xi_H^t \times \lambda_{hhH}$ from the m_{hh} distribution, as anticipated to be measured at the HL-LHC. We employ a neural network that has been trained on "realistic" m_{hh} distributions, i.e. taking into account smearing, binning, and, most crucially, Poisson noise to reflect statistical fluctuations in future measurements. This approach allows us to simulate the expected conditions under which experimental data would be obtained, making our analysis as close to reality as possible.

We employ a simple one layer network using `PyTorch` [68]. The network architecture consists of an initial batch normalization layer, followed by a single hidden layer with 64 neurons and ReLU activation, and a final output neuron with a linear activation to predict the value of $\xi_H^t \times \lambda_{hhH}$. We use the Adam optimizer (learning rate 10^{-4}, no weight decay or regularization) with mean squared error (MSE) loss. It is important to emphasize that this neural network is almost trivially small. The key point here is that this is not an application of deep learning and associated concepts, but instead a demonstration that a universal function approximator like a neural network can be used to replace classical statistical methods. In our case, we show that the network can learn to predict a value from simulated experimental data with a precision exceeding that of classical methods. More complex architectures were tested but offered no significant performance advantage, despite requiring substantially more computational resources.

As input, the model receives 16-bin histograms of the smeared m_{hh} distribution, with each bin representing a width of 50 GeV. First, these m_{hh} distributions as obtained from the theory calculation are used for training. In a later stage, these inputs are modified by applying Poisson noise, mimicking statistical uncertainties on the binned experimental data. The batch normalization layer is used to automatically rescale the count values in each bin to values close to 0 in order to speed up convergence in training. The training dataset consists of $\mathcal{O}(500-4000)$ distributions, depending on the scan settings of the 2HDM parameter space (i.e. whether we fix m_{12}^2 and m_H or not). 25% of the data is optionally retained for validation, depending on the configuration (we retain it only if we do not apply Poisson noise for training). Each model is trained for a total of 2^{15} epochs, during which a new noisy version of the input is generated at each epoch. In figure 9 we attach a schematic diagram of the workflow, specifying the input and output in each step for the different codes that were used, as well as a diagram of the NN architecture.

It is worth to note that in our approach, the NN has been trained completely on parameter points of the 2HDM. However, we expect similar results from models where the di-Higgs production process is resonantly enhanced by the contribution of an additional heavy scalar. On the other hand, during the HL run of the LHC, the searches for BSM Higgs bosons will continue. We assume here a potentially realistic situation, in which the heavy $\mathcal{CP}$-even Higgs will have been discovered and its mass will have been measured with a sufficiently high accuracy (see below). Furthermore, the rest of the spectrum, the $\mathcal{CP}$-odd Higgs boson, as well as the pair of charged Higgs bosons will have been found (and their masses will have been determined, which play no role in our analysis, but could influence the theoretical prediction of m_{hh} around the resonance at higher orders). In the case that in the future a different scalar sector will be detected experimentally, the NN analysis performed here will have to

JHEP06(2026)274

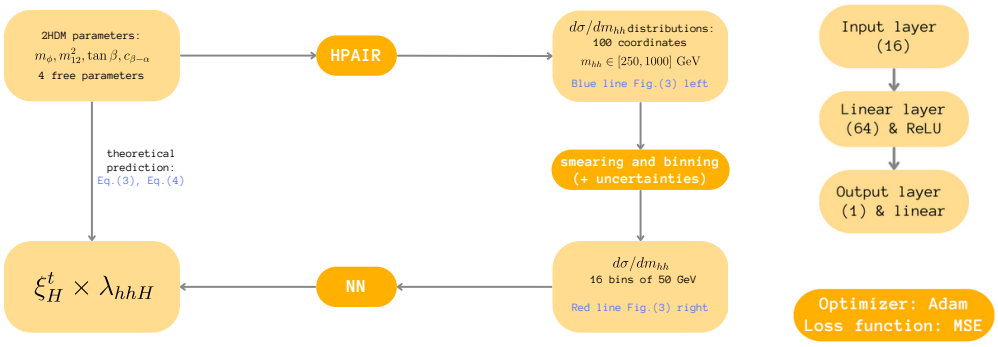

Figure 9. Workflow (left) and NN architecture (right) (see text for details).

be adopted to the then favored Higgs-boson sector. While all this at the current time is a hypothetical scenario, we use this as a viable example to demonstrate the extraction of $\xi_H^t \times \lambda_{hhH}$ from the di-Higgs boson measurements at the HL-LHC.

5 Results

In this section, we will present the results of our analysis, starting from relatively simple scenarios to more and more realistic ones by taking into account the various anticipated experimental uncertainties. We train the NN and obtain the NN predictions as described in section 4, if not specified differently.

5.1 Variation of m_{12}^2 vs. fixed m_{12}^2

We start our analysis with the effects of a variation of m_{12}^2, which is difficult to measure experimentally, see ref. [56] for some proposal.

In figure 10 we show the parameter scan of the benchmark scenario defined in section 2.3 projected in two-dimensional planes of $t_\beta - c_{\beta-\alpha}$ (left), $m_{12}^2 - c_{\beta-\alpha}$ (middle) and $m_{12}^2 - t_\beta$ (right). The upper row shows the projections with m_{12}^2 restricted to values that obey $m_{12}^2 = m_H^2 \cos\alpha^2/t_\beta$ according to eq. (2.6). This condition, however, was artificially introduced to facilitate finding a parameter space in agreement with the theoretical constraints. In order to demonstrate the effects of this artificial choice, in the lower row of figure 10 we show the three projections including all m_{12}^2 values that yield agreement with the theoretical constraints. This results in a truly 3-dimensional scan in the three remaining free parameters (once m_ϕ is fixed), i.e. $c_{\beta-\alpha}$, t_β, and m_{12}^2. In the lower right panel one can observe that for large t_β values the allowed region for m_{12}^2 is constrained to the one that fulfills eq. (2.6), while for low t_β the value of m_{12}^2 remains relatively unconstrained. The color coding indicates the values of $\xi_H^t \times \lambda_{hhH}$ and the white regions are disallowed by either of the considered constraints.

We study the effect of not fixing the value of m_{12}^2 by applying the NN on the scenario with fixed m_{12}^2 in comparison with the case of varied m_{12}^2. In figure 11 we show the result of the prediction for $\xi_H^t \times \lambda_{hhH}$ when the model is trained on the benchmark plane fixing m_{12}^2 according to eq. (2.6) on the left and without any assumption on the right. For these plots, we have trained the network on 75% of the points of the corresponding scenario (fixed m_{12}^2

JHEP06(2026)274

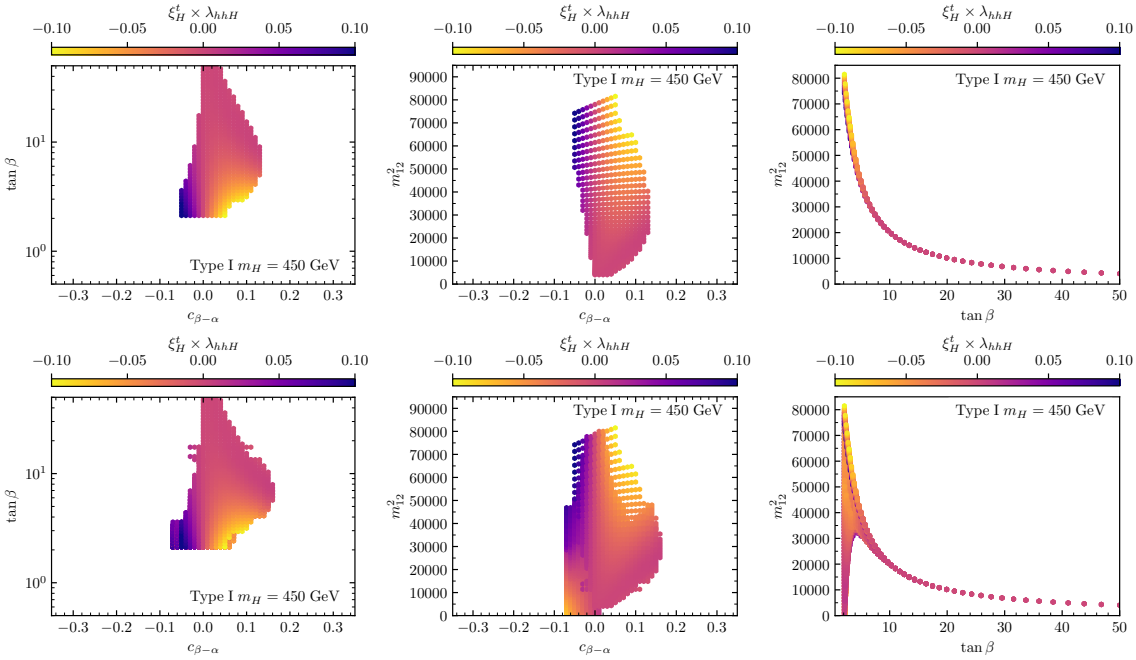

Figure 10. Upper row: $\xi_H^t \times \lambda_{hhH}$ in the allowed region of the original benchmark plane with $m_H = m_{H^\pm} = m_A = 450\,\text{GeV}$ and $m_{12}^2 = m_H^2 \cos\alpha^2/t_\beta$, projected in the three possible dimensions → planes. Lower row: same of the above → Same as upper row, but for all theoretically allowed values of m_{12}^2 (see text).

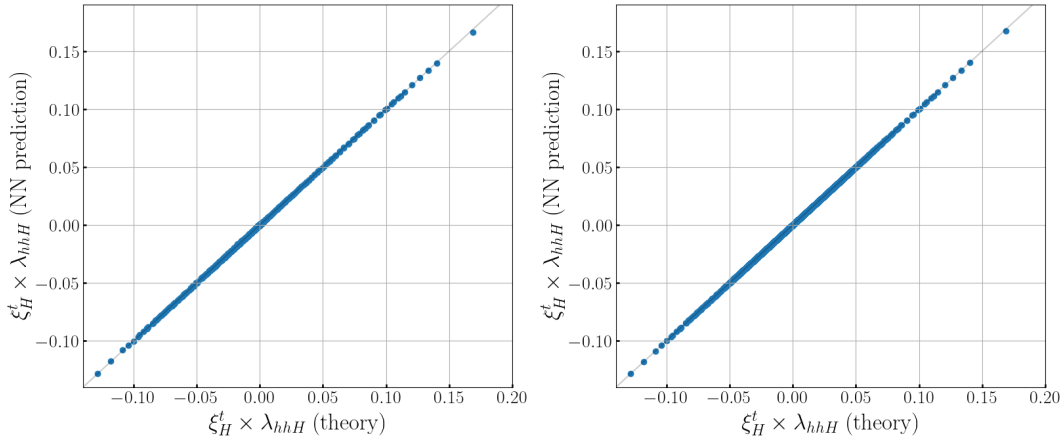

Figure 11. Prediction of $\xi_H^t \times \lambda_{hhH}$ in the original benchmark plane (left) and the three dimensional benchmark plane with $m_\phi = 450\,\text{GeV}$ and t_β, $c_{\beta-\alpha}$, and m_{12}^2 as free parameters (right).

on the left and free m_{12}^2 on the right) and tested it on the other 25% of the points. Using a 4-fold split of the full data (thus evaluating $\xi_H^t \times \lambda_{hhH}$ for the whole allowed parameter space) we observe that the prediction of the parameter $\xi_H^t \times \lambda_{hhH}$ by the NN over the whole plane is perfect for both scenarios. Thus, it can be seen that the effect of varying m_{12}^2 does not have a relevant effect on our analysis. Consequently, for the sake of computational efficiency we will continue the following analysis with the model based on the benchmark plane described in section 2.3, i.e. with m_{12}^2 fixed according to eq. (2.6).

JHEP06(2026)274

5.2 Variation of m_H

In order to investigate the NN results on the anticipated precision of m_H we analyzed the results in four different "models". The NN is trained on

`Model I`: only the example benchmark plane in figure 2 is used for the training.

`Model II`: `Model I` plus two planes with $m_H = 450 \pm 1\,\text{GeV}$ and two planes with $m_H = 450 \pm 5\,\text{GeV}$.

`Model III`: `Model II` plus two planes with $m_H = 450 \pm 10\,\text{GeV}$.

`Model IV`: `Model III` plus two planes with $m_H = 450 \pm 15\,\text{GeV}$.

While in the previous subsection the nominal benchmark value of $m_H = 450\,\text{GeV}$ has been used, we now assume that the "real" value differs by $\pm 7\,\text{GeV}$ from the measured one,[6] resulting in a corresponding distortion of the m_{hh} distributions. The results are shown in figure 12 for the four models defined above: the NN output of `Model-I, -II, -III, -IV` is presented in the upper left, upper right, lower left and lower right plot, respectively. The blue (orange) points show the prediction of the NN for two distinct test cases: $m_H = 443\,(457)\,\text{GeV}$. In both these cases we show the values of $\xi_H^t \times \lambda_{hhH}$ predicted by the NN for the whole set of allowed parameter points in the $c_{\beta-\alpha}$–t_β plane. It can be observed that the NN prediction improves significantly once we take into account as many datasets, i.e. a large variation of m_H, as possible in order to train the NN. We conclude that a possible uncertainty in m_H can be brought easily under control by a sufficient NN training for different m_H values, in particular we find that the dataset needs to contain points with an uncertainty twice as large as the measured values (e.g. for a $\pm 7\,\text{GeV}$ uncertainty in m_H, the dataset needs information from $\pm\,15\,\text{GeV}$ deviation in m_H distributions). Consequently, in the following we will either keep m_H fixed to its nominal value of $m_H = 450\,\text{GeV}$, or use `Model IV`, when the uncertainty on m_H is taken into account.

5.3 Experimental uncertainties in the m_{hh} determination: final results for the NN determination of $\xi_H^t \times \lambda_{hhH}$

So far we have assumed that the experimental result for the measured m_{hh} distribution agrees exactly with the theoretical prediction. However, a realistic distribution is subject to statistical and systematic uncertainties in the experimental measurement of m_{hh}, as discussed in section 2.6. While an assessment of the systematic uncertainties goes far beyond the scope of this work, for the final results we will include the anticipated statistical uncertainties in the m_{hh} measurement. In order to take this into account, we repeat our analysis for the three cases discussed so far: 1) the initial benchmark scenario with m_H and m_{12}^2 fixed, 2) relaxing the constraint on m_{12}^2, and 3) relaxing the constraint on m_H.[7]

Our final NN analysis was performed as follows. The NN was trained on the full set of `Model I`. For each point in the plane we derived 2^{15} statistically smeared m_{hh} distributions,

[6] The experimental precision on BSM Higgs-boson masses at the LHC remains to be analyzed in detail; see, however, ref. [69].

[7] Relaxing both constraints at the same time would lead to a very large increase in the data points that would be computationally challenging.

JHEP06(2026)274

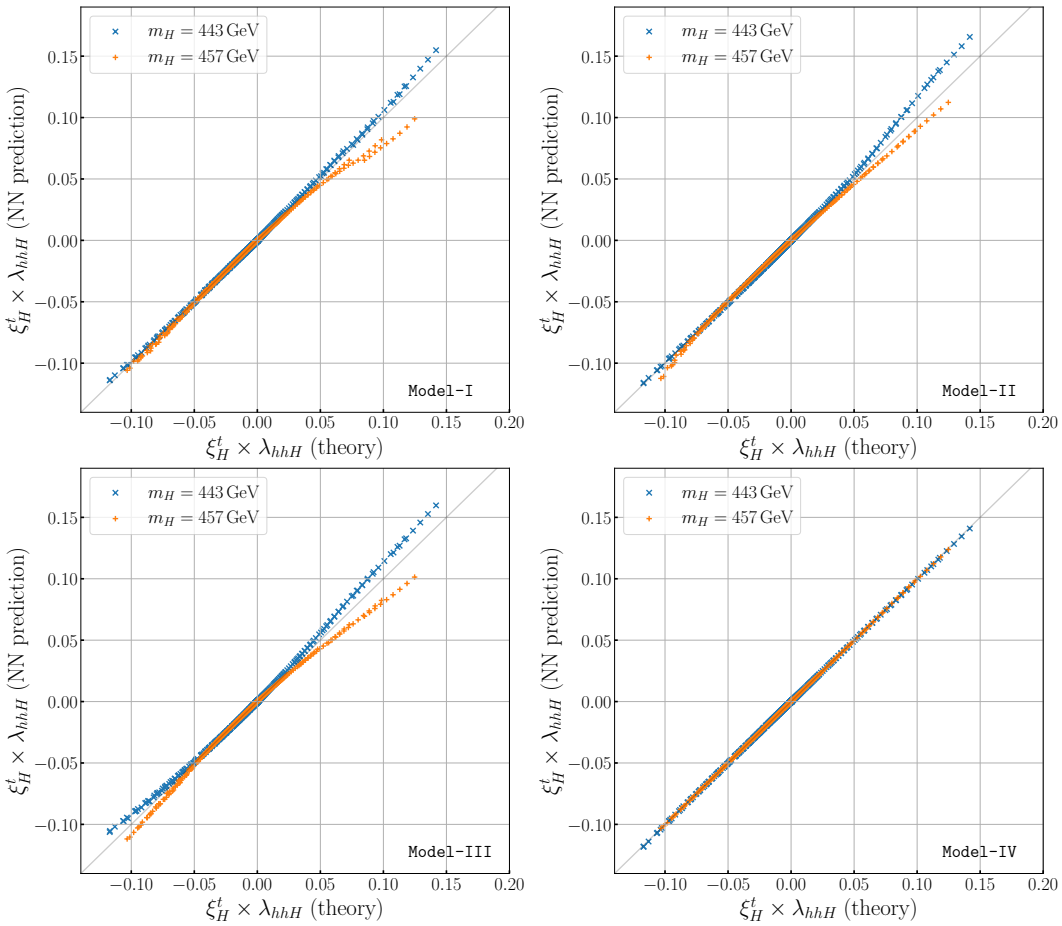

Figure 12. Results of our NN analysis for $\xi_H^t \times \lambda_{hhH}$ in `Model-I, -II, -III, -IV` in the upper left, upper right, lower left and lower right plot, respectively. The blue (orange) points show the results for $m_H = 443(457)\,\text{GeV}$ (see text). The gray line is there to guide the eye and corresponds to the coincidence of the values obtained from the NN with the values predicted by theory.

for which we use the Poisson distribution (see section 2.6 for a definition of the smeared, and section 3 for the related smearing of the data). We then obtained the $\xi_H^t \times \lambda_{hhH}$(NN predicted) values for 256 test m_{hh} distributions that were also statistically smeared with their respective Poisson distribution. The result is shown in the form of a scatter plot as a function of the true $\xi_H^t \times \lambda_{hhH}$(theory) value in the left plot of figure 13. (The points marked in red and green in this plot are going to be discussed in more detail further below.) In comparison with the statistically unsmeared case, see the left plot in figure 11, a substantial broadening of the NN predicted $\xi_H^t \times \lambda_{hhH}$ values can be observed, resulting in an apparently strong degradation of the NN capability to find the correct value of $\xi_H^t \times \lambda_{hhH}$. This, however, neglects the fact that most of the points that have a very different $\xi_H^t \times \lambda_{hhH}$(NN predicted) value w.r.t. the corresponding $\xi_H^t \times \lambda_{hhH}$(theory) value are "outliers", and most of the predicted results obtained are still close to the diagonal. In order to demonstrate this, on the right plot we present the density distribution of the points of the plot on the left, where we have normalized to 100 the density at each bin of $\xi_H^t \times \lambda_{hhH}$(theory). Thus, the density in one vertical strip corresponding to a given value of $\xi_H^t \times \lambda_{hhH}$(theory) adds up to 100. One

JHEP06(2026)274

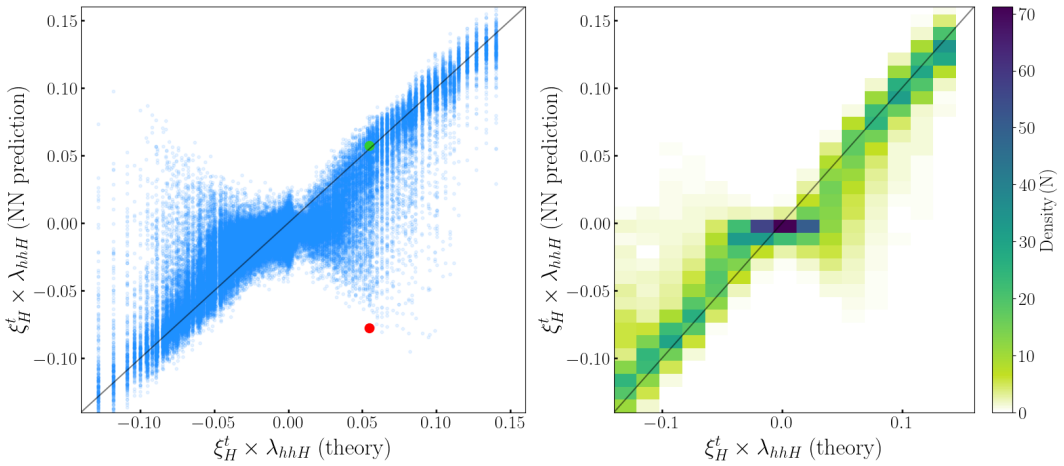

Figure 13. Prediction of $\xi_H^t \times \lambda_{hhH}$ for points with $m_H = 450\,\text{GeV}$ with the m_{hh} distributions statistically smeared according to their respective Poisson distributions. We show the results in a scatter plot (left) and a density plot (right), where in the latter case the color represents the density of points that falls into this grid. The points marked in green and red in the left plot are example points discussed in the text.

can see that the highest density of points in the prediction falls into the diagonal line of $\xi_H^t \times \lambda_{hhH}(\text{NN predicted}) = \xi_H^t \times \lambda_{hhH}(\text{theory})$. This means that despite there are some outliers, the NN prediction is reasonably accurate for the majority of points.

We have conducted a more detailed analysis of two distinct smeared distributions, both corresponding to the same benchmark point indicated by a cyan cross in figure 2. Their predictions are represented in the left panel of figure 13, with the green dot ("close"), which lies near the theoretical value, and the red dot ("far"), which deviates significantly from the theoretical prediction. The parameters of the benchmark point are given in table 1, together with the NN predictions of the two smeared distributions. We show these distributions in figure 14, where the blue m_{hh} distribution is identical in both plots and has been calculated with $\xi_H^t \times \lambda_{hhH}(\text{theory})$. As before, the vertical bars indicate the statistical uncertainty for each bin. Based on these uncertainties, the green (left plot) and red (right plot) m_{hh} distributions are the ones that were randomly chosen for this point to be fed to the trained NN to evaluate $\xi_H^t \times \lambda_{hhH}(\text{NN predicted})$. In the left plot we plot the distribution corresponding to the "close" point, which has a $\xi_H^t \times \lambda_{hhH}(\text{NN predicted})$ value close to $\xi_H^t \times \lambda_{hhH}(\text{theory})$, see table 1, and which is represented by the green point in the left plot of figure 13. It can be seen that the key to get this good agreement in the NN prediction is the small deviation of the randomly chosen distribution from the one based on $\xi_H^t \times \lambda_{hhH}(\text{theory})$, particularly in the five bins around the resonance at $m_{hh} = 450\,\text{GeV}$. This argument becomes clearer when looking at the right plot, corresponding to the "far" point. In this case the five bins closer to the resonance exhibit a shift in the relative height of the true distribution w.r.t. the randomly chosen one, based on the statistical uncertainties. These shifts produce event numbers with three out of five outside the respective $1\,\sigma$ interval of the respective bin, resulting in the large difference of $\xi_H^t \times \lambda_{hhH}(\text{NN predicted})$ and $\xi_H^t \times \lambda_{hhH}(\text{theory})$. In this case, the NN even predicts the wrong sign, indicated by the red point in the left plot of figure 13.

JHEP06(2026)274

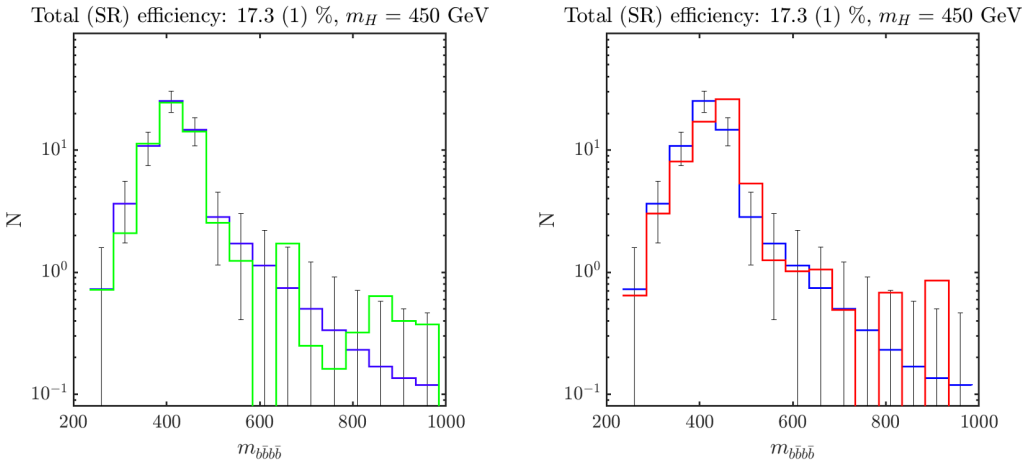

Figure 14. Examples of invariant mass distributions of the point marked by a cyan cross in figure 2 with the parameters given in table 1. The blue distributions are based on $\xi_H^t \times \lambda_{hhH}$(theory), and the vertical bars represent the statistical uncertainties. The green (left) and red (right) distributions are randomly chosen based on these uncertainties (see text).

t_β	$c_{\beta-\alpha}$ $(s_{\beta-\alpha} > 0)$	$\xi_H^t \times \lambda_{hhH}$(th)	$\xi_H^t \times \lambda_{hhH}$(NN) "close"	$\xi_H^t \times \lambda_{hhH}$(NN) "far"
4.38	-0.04	0.0547	0.0572	-0.0775

Table 1. Parameters for the point marked by the cyan cross in figure 2. The value of $\xi_H^t \times \lambda_{hhH}$(NN predicted) "close" ("far") corresponds to the random distribution shown in green (red) in the left (right) plot of figure 14. ($\xi_H^t \times \lambda_{hhH}$(th) is short for $\xi_H^t \times \lambda_{hhH}$(theory), $\xi_H^t \times \lambda_{hhH}$(NN) is short for $\xi_H^t \times \lambda_{hhH}$(NN predicted).) The "close" point is marked in green in figure 13, and the "far" point is marked in red.

This raises the question whether the NN prediction of $\xi_H^t \times \lambda_{hhH}$ can be improved by increasing the number of "strongly" statistically Poisson smeared distributions. To analyze this, we trained the network on data with a larger uncertainty range (allowing 2σ error bars with the means of the Poisson distribution at the theoretical values of $\xi_H^t \times \lambda_{hhH}$) and validated it on data with a 1σ uncertainty. In this way we provide to the net more information about the points that deviate the most from the prediction, so that it can *learn* more about them. The results can be seen in figure 15, where one can observe that the training extended to more strongly statistically smeared distributions visibly improves the result. (A more quantitative analysis will be given below.) We also tested whether a further increase of the statistical uncertainty in the training sample would further improve the NN prediction of $\xi_H^t \times \lambda_{hhH}$. However, we found that the increase of the Poisson uncertainty by a factor of ~ 2 is close to the optimum value (see our more quantitative analysis below).

As a further step, we also took into account the effect of statistical smearing on the dataset with different values of m_{12}^2 and m_H, i.e. `Model IV`, by changing one at a time. We show the results of each of these steps in figure 16, where the upper row shows the effect of the uncertainty in the value of m_{12}^2 and the lower row shows the effect of the uncertainty in m_H by including the statistically smeared samples in the NN training with the 3D dataset in the former case and the `Model IV` in the latter. Again the results are shown for 256 points

JHEP06(2026)274

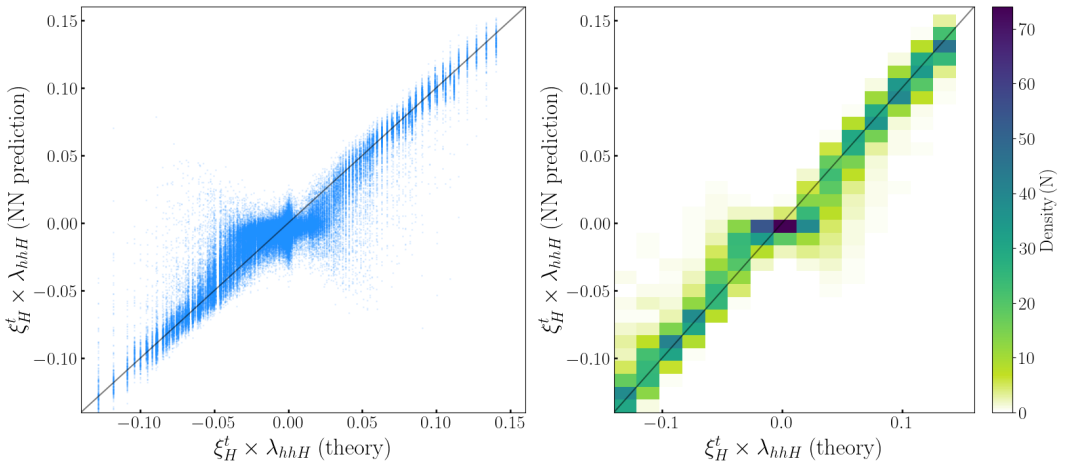

Figure 15. Prediction of $\xi_H^t \times \lambda_{hhH}$ for points with $m_H = 450\,\text{GeV}$ with the m_{hh} distributions statistically smeared according to twice their respective Poisson distributions. The color coding is as in figure 13.

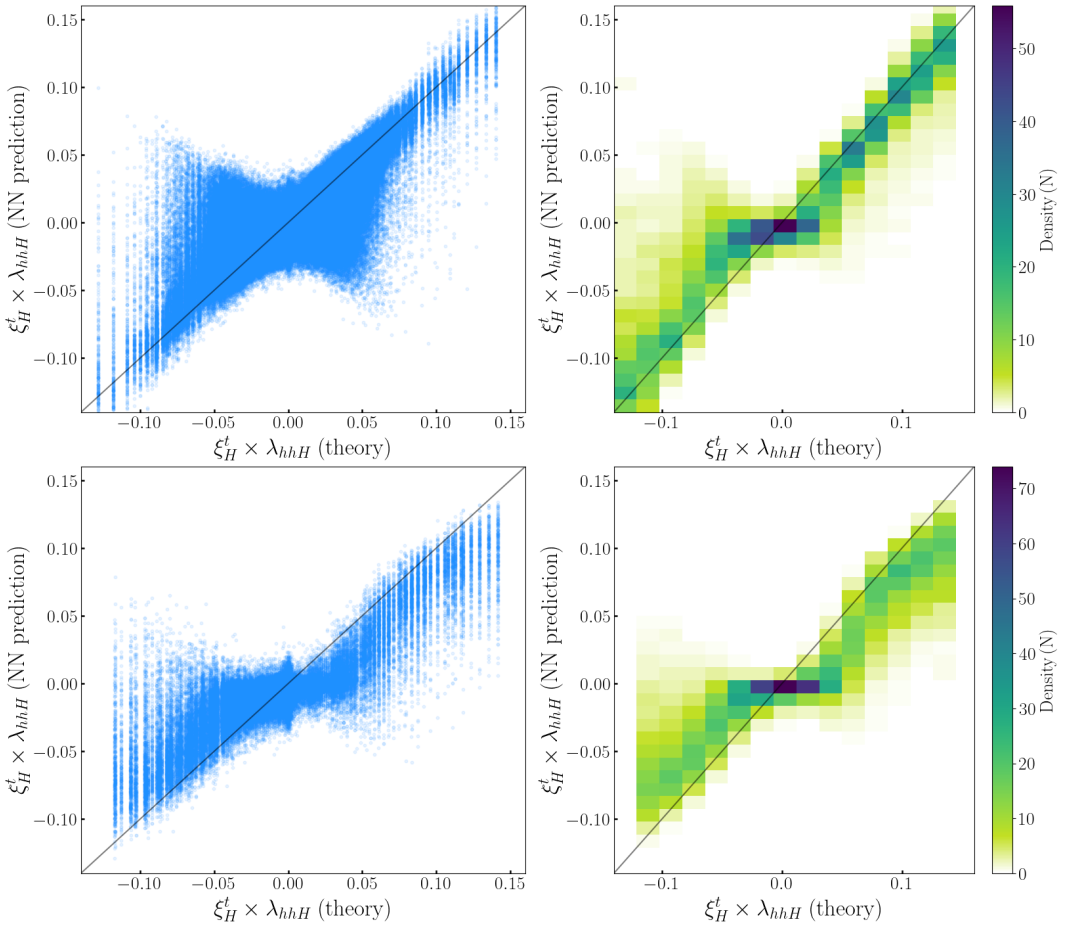

Figure 16. Predictions for statistically smeared data. Upper row for free m_{12}^2 and lower row for $m_H = 443\,\text{GeV}$ (as a test case) employing Model IV.

JHEP06(2026)274

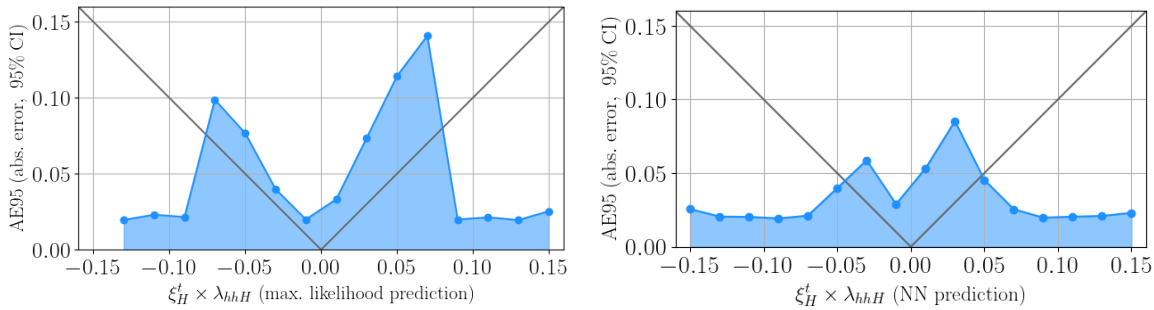

Figure 17. Comparison of the AE95 with MLE (left) and NN (right) depending on the value of $\xi_H^t \times \lambda_{hhH}$.

in the scatter plot on the left side of the figure, and the corresponding density of points is colored on the right side. Including the uncertainties related to the values of m_{12}^2 and m_H clearly enlarges the span of the predicted values by the NN for a given value of $\xi_H^t \times \lambda_{hhH}$. However, the corresponding density plots still display a sufficiently good precision, which will be further discussed in the next subsection.

Our conclusion from this analysis is that the determination of the variable $\xi_H^t \times \lambda_{hhH}$ with the correctly trained NN is possible, even taking into account all relevant experimental uncertainties (where we did not take into account the — to us unknown — systematic experimental uncertainties). For this determination it is crucial to reduce the statistical uncertainties, and the result depends strongly on the experimental efficiencies. But even with the low acceptance of $\epsilon = 0.17 \times 0.01$ the right plot of figure 13 demonstrates that an experimental determination of $\xi_H^t \times \lambda_{hhH}$ is possible.

5.4 Comparison of NN and statistical analyses

In this section, we compare the improvement in the performance of the NN with respect to the classical MLE method on the original benchmark scenario. We furthermore give more quantitative results in the comparison of the results with different assumptions on the experimental uncertainties. To quantify this, we have used the 95% confidence interval for the absolute error of $\xi_H^t \times \lambda_{hhH}$, AE95. We trained the NN and the MLE on an equal number of samples. The interpretation of AE95 is that the error in the prediction of $\xi_H^t \times \lambda_{hhH}$ is at 95% CL smaller than AE95, therefore, the smaller the value of AE95, the more reliable the prediction of the corresponding method.

We find that the error is highly sensitive to the value of $\xi_H^t \times \lambda_{hhH}$ as illustrated in figure 17. The left plot shows the AE95 for $\xi_H^t \times \lambda_{hhH}$ computed using the MLE method, while the right plot displays the results from the NN. Both approaches reveal a similar overall trend: the largest errors occur near $\xi_H^t \times \lambda_{hhH} = \pm 0.05$ when using the NN (± 0.07 using the MLE method) and decrease for larger absolute values as well as for smaller ones. Additionally, the minimum error appears near zero, with a slight preference for negative values.

We can define the ranges where the 95% CL error is larger than the value itself as the crossing points between the blue and gray lines in figure 17, this comparison yields the

JHEP06(2026)274

$\xi_H^t \times \lambda_{hhH}$ intervals:

$$\begin{aligned} \xi_H^t \times \lambda_{hhH} < -0.0758 \quad \text{or} \quad \xi_H^t \times \lambda_{hhH} > 0.0799 \quad \text{for MLE,} \\ \xi_H^t \times \lambda_{hhH} < -0.0459 \quad \text{or} \quad \xi_H^t \times \lambda_{hhH} > 0.0450 \quad \text{for NN } (1\sigma). \end{aligned} \tag{5.1}$$

These intervals can be interpreted as the sensitivity ranges of $\xi_H^t \times \lambda_{hhH}$ with each method. It is clearly visible that the NN method yields a larger sensitivity to the BSM couplings.[8]

We can also use these intervals to compare the regions of the original plane that can be probed with both methods. This region is shown in figure 18. In this plot we display the original benchmark plane with the color coding indicating the magnitude of $\xi_H^t \times \lambda_{hhH}$. Inside the solid contour lines the determination of $\xi_H^t \times \lambda_{hhH}$ would not be conclusive with each respective method. We point out several observations: 1) The maximum likelihood (red contour) is worse than the NN (cyan contour) for the determination of $\xi_H^t \times \lambda_{hhH}$. 2) The p-value, obtained with the toy-MC approach discussed in section 3, (dashed black contour) cannot perform a parameter estimation. It can only exclude the null hypothesis of the SM, and in this respect it is competitive with the NN. 3) The NN can do both the hypothesis test and the parameter estimation simultaneously. While being comparable with the classical methods in the first case (hypothesis testing), it outperforms them in the second case (parameter estimation). We have furthermore obtained the AE95 precision for the case that the Poisson uncertainty has been doubled for the training samples. In this case the AE95 limits for $\xi_H^t \times \lambda_{hhH}$ are,

$$\xi_H^t \times \lambda_{hhH} < -0.0334 \quad \text{or} \quad \xi_H^t \times \lambda_{hhH} > 0.0341 \quad \text{for NN } (2\sigma). \tag{5.2}$$

shown with a solid blue line in figure 18, the blue line for positive $\xi_H^t \times \lambda_{hhH}$lies below the black line. The results are better than with other methods, in particular for negative values of $\xi_H^t \times \lambda_{hhH}$. Training the NN with points at 2σ from the mean value, provides more information about the points that deviate the most from the prediction, so that it can learn more about them. This explains how the NN becomes more robust. The optimal range (around 2σ) means that there is a balance: too little noise limits learning from outliers, while too much noise makes it difficult for the NN to learn meaningful patterns. We have verified explicitly that training the net on values with even larger uncertainties worsens the results.

We therefore conclude that the NN is not only the best method for the BSM parameter determination, as it can do both hypothesis testing and parameter estimation in a more efficient way, it also has the potential to be improved with a dedicated analysis of the data.

Using the AE95 metric, we can also compare the performance of the net for the different training sets of data, taking into account various possible future experimental uncertainties. In this case the one corresponding to `Model I` and shown in figure 13 is called dataset (1), the one including the m_{12}^2 as a free parameter shown in the upper row of figure 16 is named

[8]We note here that taking the efficiencies corresponding to at least 4 b-tagged jets (see the discussion in section 2.6) weakens the 95% CL interval for the determination of $\xi_H^t \times \lambda_{hhH}$ by roughly a factor of two. However, the qualitative conclusions and the NN improvement over the classical method remain unchanged. We stick to the efficiencies corresponding to at least two b-tagged jets because the HL-LHC tagging efficiencies and final analysis strategies are not yet fixed, and potential improvements in detector performance and/or effective integrated luminosity may partially compensate for the reduced efficiency.

JHEP06(2026)274

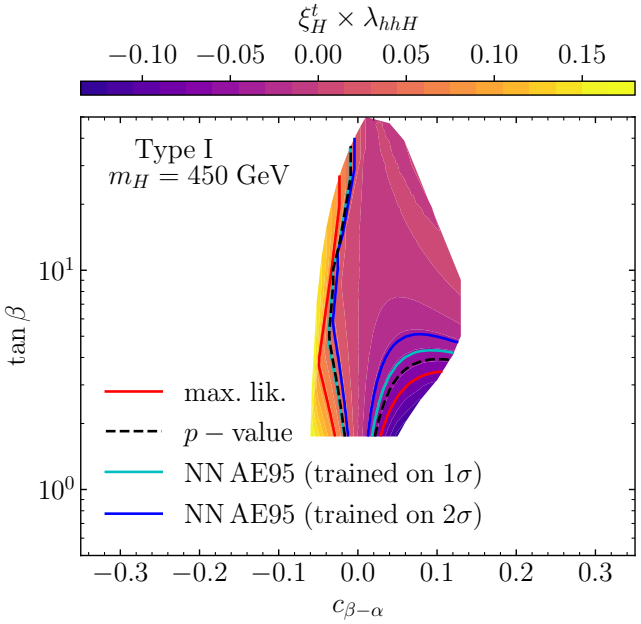

Figure 18. Comparison between the NN and the classical MLE approach.

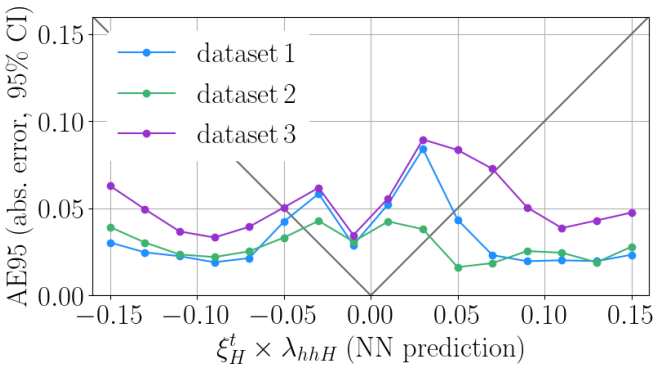

Figure 19. AE95 for the different datasets employed. Dataset 1 refers to the original benchmark plane, dataset 2 to the plane with m_{12}^2 free and dataset 3 is the one with the uncertainty in the value of m_H.

dataset (2), and the one shown in the lower row of figure 16, the one corresponding to `Model IV`, we call dataset (3). We obtain the following results

$$\begin{aligned}
&\xi_H^t \times \lambda_{hhH} < -0.0459 \quad \text{or} \quad \xi_H^t \times \lambda_{hhH} > 0.0450 \quad \text{for dataset (1)},\\
&\xi_H^t \times \lambda_{hhH} < -0.0386 \quad \text{or} \quad \xi_H^t \times \lambda_{hhH} > 0.0338 \quad \text{for dataset (2)},\\
&\xi_H^t \times \lambda_{hhH} < -0.0503 \quad \text{or} \quad \xi_H^t \times \lambda_{hhH} > 0.0712 \quad \text{for dataset (3)}.
\end{aligned} \tag{5.3}$$

The respective AE95 for each value $\xi_H^t \times \lambda_{hhH}$ are shown in figure 19. The key features are that interestingly, the net performs better for the dataset (2) than for dataset (1), i.e. adding more free parameters of the model to the dataset improves its performance. It performs worse for the dataset (3) than for the dataset (1), specially for positive $\xi_H^t \times \lambda_{hhH}$ values,

JHEP06(2026)274

which suggests that an uncertainty in the measurement might be a source of error in the NN approach, which, however, can be brought under control by including a larger training sample with more m_H values. We note in particular that the results for the dataset (2) appear to be better than the ones for dataset (3). The reason is that having m_{12}^2 as a free parameter enlarges the learning sample for the NN for values of $\xi_H^t \times \lambda_{hhH}$ close to the alignment limit, which are the ones for which the prediction is more challenging. In contrast, the values of $\xi_H^t \times \lambda_{hhH}$ further away from the alignment limit are easier to distinguish from the SM and these are present in the original scenario, as this is the one that allows largest mass splittings between the couplings while being consistent with the perturbative unitarity constraints. By changing m_{12}^2 we deviate from the optimal region for perturbative unitarity, and larger deviations in $\xi_H^t \times \lambda_{hhH}$ from the SM become disallowed. We conclude that a larger sample, even with more free parameters, would improve the determination of the $\xi_H^t \times \lambda_{hhH}$ by the NN.

5.5 Future prospects

The results in the previous subsection have been obtained using the experimental efficiencies as given by ATLAS [66]. However, it is conceivable that the efficiencies during the HL run of the LHC might improve due to a better knowledge of the systematic uncertainties with increasing luminosity. As a hypothetical scenario we analyze here the level of improvement in the NN determination of $\xi_H^t \times \lambda_{hhH}$ in the case that each of the two efficiencies, ϵ_{TOT} and ϵ_{SR}, are improved by a factor of two respectively. This increase of the number of events by a factor of four would yield a factor of 1/2 in the statistical uncertainty for each bin in the m_{hh} distribution.

In figure 20 we show the results of our NN analysis assuming this improvement of the combined efficiency by a factor of four, i.e. the same as in section 5.3, but with relative statistical uncertainties smaller by a factor of 1/2. We evaluate the improvements on the initial benchmark plane (first row), on the plane with the free m_{12}^2 (middle row) and, additionally, the free m_H (lower row). We show the scatter plots on the left and the density plots on the right. All the scatter plots in the left side of figure 20 show a substantial improvement compared to the corresponding (left) plots in figures 13 and 16. The improvement becomes even more apparent in the comparison of the density plots on the right of figure 20 with the corresponding (right) plots of figures 13 and 16. With the efficiencies improved by a factor of two a NN determination of $\xi_H^t \times \lambda_{hhH}$ at the level of 10–20% appears to be feasible.

6 Conclusions

Di-Higgs production at the (HL-)LHC provides the unique opportunity to measure triple Higgs couplings to assess the form of the Higgs potential. Beyond-SM Higgs sectors, as suggested by the open questions of the SM, not only introduce new Higgs bosons into the model, but also the corresponding BSM THCs. Existing LHC studies of THCs concentrated on the THC of the Higgs boson at $\sim 125\,\text{GeV}$, λ_{hhh}. A second ($\mathcal{CP}$-even) Higgs boson in the spectrum can contribute via an s-channel exchange to the process $gg \to hh$ and interfere with the s-channel exchange of the h and the di-Higgs production via a top-quark box. This contribution is $\propto \xi_H^t \times \lambda_{hhH}$, the product of the top Yukawa coupling of the new Higgs boson and its THC with two 125 GeV Higgs bosons.

JHEP06(2026)274

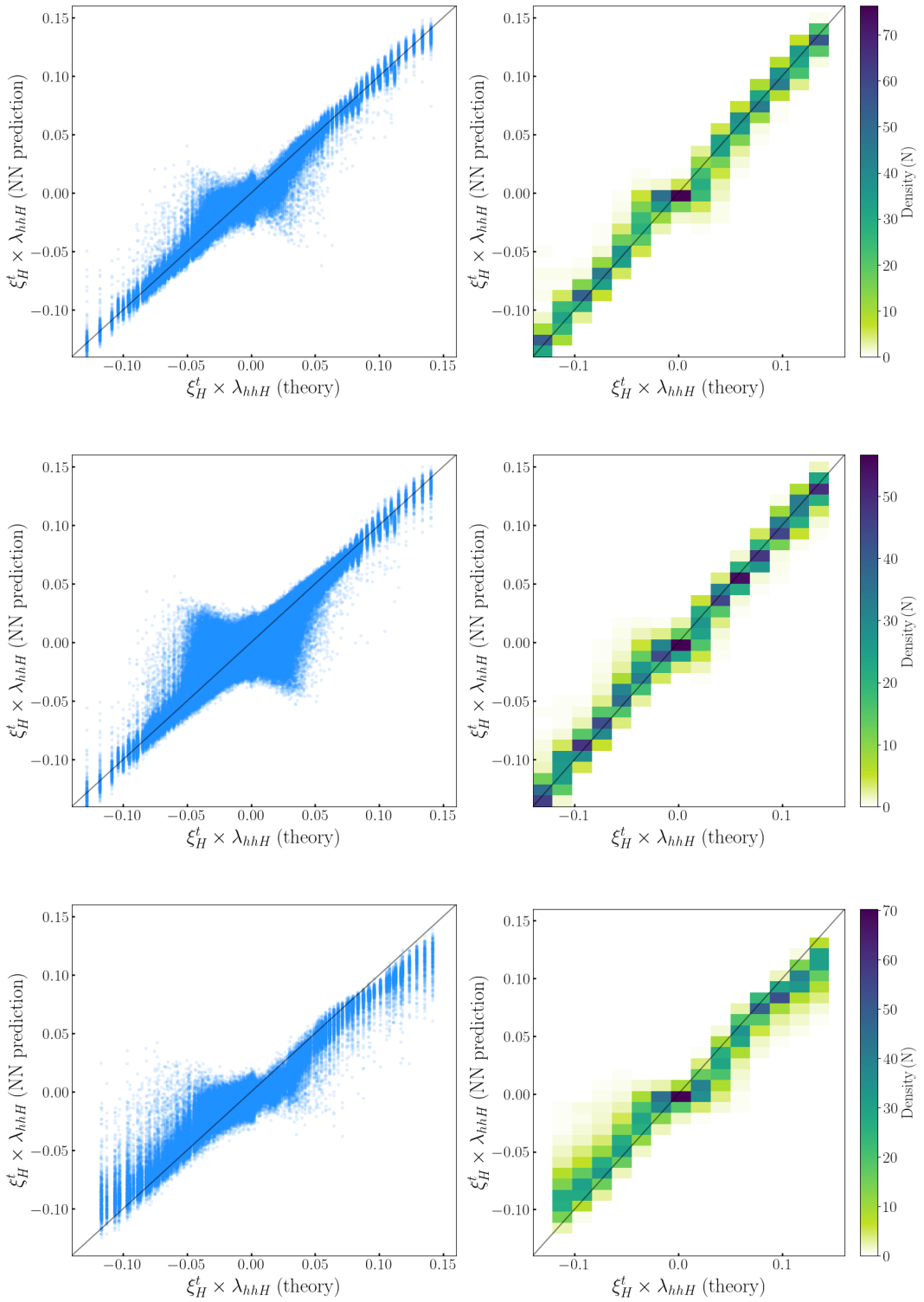

Figure 20. Same as figure 13 but with a factor 4 improvement in the total efficiency ϵ (see text).

The main goal of this work is to assess how well a NN can determine $\xi_H^t \times \lambda_{hhH}$ based on a "realistic" measurement of the m_{hh} distributions from the process $gg \to hh$ at the HL-LHC. With "realistic" we refer to a 15% smearing in m_{hh} (corresponding to the

JHEP06(2026)274

anticipated experimental accuracy), as well as to a 50 GeV binning in m_{hh} (corresponding to the anticipated finite energy resolution). As a concrete example, we focused on the 2HDM, providing naturally a second $\mathcal{CP}$-even Higgs boson, H. We assumed a possible future scenario in which the H will have been discovered with $m_H \approx 450\,\text{GeV}$, but the soft $\mathbb{Z}_2$ symmetry-breaking parameter m_{12}^2 is undetermined. The two mixing angles, parametrized by t_β and $c_{\beta-\alpha}$, are restricted by theoretical constraints, but also significantly by the LHC Higgs-boson rate measurements and the BSM Higgs-boson searches. These two mixing angles constitute the two main free parameters of our analysis.

For our analysis we employ a NN with the following characteristics: an input layer with 16 nodes (corresponding the input nodes for the binned m_{hh} distribution of a certain $\xi_H^t \times \lambda_{hhH}$ value) is followed by one layer with 64 nodes. The final layer consists of one neuron that predicts the value of $\xi_H^t \times \lambda_{hhH}$. In general, the NN is trained on a subset of m_{hh} distributions and then used to "predict" $\xi_H^t \times \lambda_{hhH}$ corresponding to the distributions it was not trained on. In a first step we demonstrated that the result for $\xi_H^t \times \lambda_{hhH}$ is not influenced by a variation in m_H (assuming a possible future HL-LHC accuracy for this mass), nor by a variation of m_{12}^2, the parameter of the 2HDM Higgs potential that is most difficult to access.

Focusing on the $c_{\beta-\alpha}$–$\tan\beta$ plane with m_H and m_{12}^2 fixed, but neglecting statistical and systematical uncertainties, we demonstrated that the NN can predict almost perfectly the value of $\xi_H^t \times \lambda_{hhH}$ from the (smeared and binned) m_{hh} input.

To properly take into account further uncertainties of the measurement, we incorporated the anticipated statistical uncertainties in m_{hh}, assuming the $b\bar{b}\,b\bar{b}$ final state and $3\,\text{ab}^{-1}$ of integrated luminosity (systematic uncertainties are beyond the scope for this exploratory theoretical study). For a given theoretical prediction of m_{hh} we evaluated the statistical uncertainty based on the expected numbers of events and the anticipated experimental efficiencies in the $b\bar{b}\,b\bar{b}$ final state. Taking these uncertainties as a Poisson 1σ uncertainty, we obtained for each point in the $c_{\beta-\alpha}$–$\tan\beta$ plane 2^{15} Poisson smeared m_{hh} distributions. For these "realistic" m_{hh} distributions the NN (trained only on the "correct" m_{hh} distributions) then predicted the corresponding $\xi_H^t \times \lambda_{hhH}$ value. We found that the highest density of $\xi_H^t \times \lambda_{hhH}$ values in the prediction falls into the diagonal line of $\xi_H^t \times \lambda_{hhH}(\text{NN predicted}) = \xi_H^t \times \lambda_{hhH}(\text{theory})$. This demonstrates that despite there are some outliers, the majority of predictions are concentrated along the diagonal. In particular, the region around the identity line shows a point density exceeding 70%, indicating that the neural network accurately captures the dominant trend between theory and prediction.

We have compared these results to the classical maximum likelihood estimation (MLE) methods for parameter estimation, and the results with NN proved to be better, with an even larger gain if the net is trained more on values that are outliers from the theoretical prediction. This not only shows the comparative improvement of machine learning versus classical statistics, it also highlights the potential for an improvement in a dedicated analysis.

Finally, we have analyzed the impact of a possible future improvement of the detector efficiencies in the HL run of the LHC. Our assumption of a factor 4 improvement in detector efficiencies would lead to a statistically significant determination of $\xi_H^t \times \lambda_{hhH}$ at the level of 10–20%. These results are subject to the future efficiencies and in particular could be worsened by a factor of two if the requirement for 4 b-tagged jets is made. On the other hand they might

JHEP06(2026)274

also improve if both ATLAS and CMS data are combined, thus yielding a larger luminosity. The overall proof of principle that a NN can determine a BSM THC and yields better results than a classical analysis, remains (largely) invariant under these assumptions and changes.

If in the future data are available that confirm the existence of a BSM scalar particle in nature, a NN with a larger accuracy can be trained, by taking into account the future mass uncertainty range and by including higher-order loop corrections to the process $gg \to hh$, and in particular to the involved THCs. We expect a more realistic and sophisticated analysis based on this future knowledge to outperform our results. Despite the fact that our analysis is model dependent, it is reasonable to assume that any model with a similar experimental signature that may then be favored by experimental data will have good prospects for the determination of the couplings involved in the resonantly produced Higgs boson pairs. In this way, NNs pave the way to determine the shape of the BSM Higgs potential.

Acknowledgments

We thank Henning Bahl for his valuable input during the early stages of this work. We thank Jana Schaarschmidt for useful insights in the statistical uncertainties of the detectors and their efficiencies. K.R. thanks Panagiotis Stylianou for useful discussions and suggestions regarding neural networks and Ana Andrade for very useful discussions regarding QCD backgrounds. The work of S.H. has received financial support from the grant PID2019-110058GB-C21 funded by MCIN/AEI/10.13039/501100011033 and by "ERDF A way of making Europe", and in part by the grant IFT Centro de Excelencia Severo Ochoa CEX2020-001007-S funded by MCIN/AEI/10.13039/501100011033. S.H. also acknowledges support from Grant PID2022-142545NB-C21 funded by MCIN/AEI/10.13039/501100011033/ FEDER, UE. The work of M.M. has been supported by the BMBF-Project 05H24VKB. K.R. acknowledges support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy — EXC 2121 "Quantum Universe" — 390833306. This work has been partially funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) — 491245950. The project that gave rise to these results received the support of a fellowship from the "la Caixa" Foundation (ID 100010434). The fellowship code is LCF/BQ/PI24/12040018.

A Impact of finite grid size of the initial m_{hh} calculations

We have used as the "theoretical prediction" the m_{hh} distributions with 100 coordinates ranging from 250 GeV to 1000 GeV, i.e. our initial bin size is 7.5 GeV. This value is precise enough for the overall m_{hh} distribution, but might be too coarse to capture well the peak around the heavy Higgs resonance. A finer grid was not used initially because of computational efficiency: we are dealing with large parameter scans and need to run HPAIR in all the cases to obtain the theoretical prediction. In order to estimate the induced error of the "large" 7.5 GeV bin size on the $\xi_H^t \times \lambda_{hhH}$ determination, which we interpret as a further source of theoretical uncertainty, we have computed ten benchmark points with an improved initial grid of 0.075 GeV. This is substantially finer than the one employed throughout this work by a factor of 100. For these points we obtained the predictions of $\xi_H^t \times \lambda_{hhH}$ with a NN trained on

JHEP06(2026)274

$c_{\beta-\alpha}$	$\xi_H^t \times \lambda_{hhH}$(theory)	$\xi_H^t \times \lambda_{hhH}$(NN predicted) (7.5 GeV bin)	$\xi_H^t \times \lambda_{hhH}$(NN predicted) (0.075 GeV bin)	Δ [%]
−0.06	0.168	0.166	0.166	<1
−0.05	0.140	0.139	0.139	<1
−0.04	0.111	0.111	0.111	<1
−0.03	0.083	0.083	0.080	4
−0.02	0.055	0.055	0.052	5
−0.01	0.027	0.027	0.024	12
0.00	0.000	0.000	0.000	<1
0.01	−0.027	−0.027	−0.025	8
0.02	−0.053	−0.053	−0.050	6
0.03	−0.079	−0.079	−0.077	3

Table 2. Predictions of the NN for m_{hh} distributions with a finer grid. Δ quantifies the numerical impact on the $\xi_H^t \times \lambda_{hhH}$(NN predicted) determination from the grid size change (see text).

the points in the benchmark plane of figure 2. The chosen points have $t_\beta = 1.74$ and are listed in table 2. One can observe that the NN prediction depends only very mildly on the initial grid size. We quantify the impact of the change of the grid size from 7.5 GeV to 0.075 GeV through

$$\Delta \equiv \frac{|[\xi_H^t \times \lambda_{hhH}(\text{NN predicted})_{0.075\,\text{GeV}} - \xi_H^t \times \lambda_{hhH}(\text{NN predicted})_{7.5\,\text{GeV}}]|}{[\xi_H^t \times \lambda_{hhH}(\text{theory})]} \,. \tag{A.1}$$

One can observe that Δ remains at the few percent level, i.e. the coarser grid size, which has been used to train the NN, yields nearly as good results as the finer one. From this we conclude that our analysis using the coarser grid size of 7.5 GeV captures all the relevant effects. Any analysis in the future based on real experimental data should, of course, use the finest grid size possible by computational efficiency.

B Details on the neural network training

In figure 21 we show the MSE loss functions and the AE95 evolutions over the training epochs for the different models. The upper row corresponds to `Model I` (dataset (1)), the middle row to `Model I` with m_{12}^2 free (dataset (2)), and the lower row to the `Model IV` (dataset (3)). These correspond to datasets with the original benchmark plane, with free m_{12}^2, and with an uncertainty in m_H, respectively.

Clearly, the training and validation losses closely track each other throughout the entire training process. There is no visible gap between them, and they decrease steadily over the epochs, which suggests that there is no overfitting. We could consider an early stopping, as the MSE does not decrease significantly after $\sim$ 20k epochs. However, we prefer to keep a longer learning rate to include further data, as each epoch learns from new Poisson smeared data.

We also mention here how we used the AE95 metric to decide on the network architecture. As shown in figure 22, there is no significant improvement in the ranges of $\xi_H^t \times \lambda_{hhH}$(NN predicted) when using a 2-hidden layer network with 128 neurons each or even a 4-hidden layer with 128 neurons each w.r.t the simple 1-hidden layer network with 64 neurons applied in the present analysis.

JHEP06(2026)274

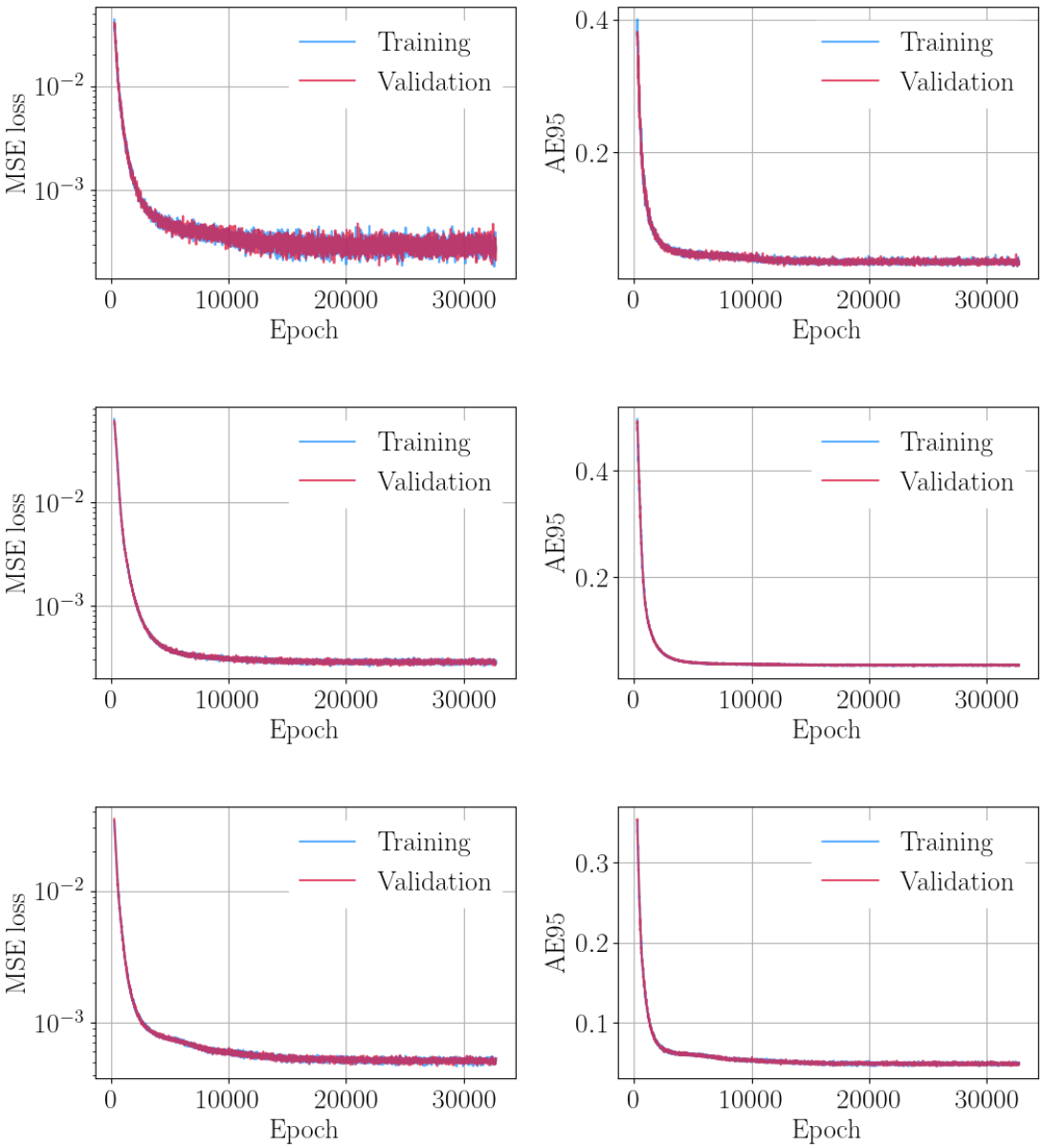

Figure 21. NN prediction of $\xi_H^t \times \lambda_{hhH}$ for the different datasets: (1) upper, (2) middle and (3) lower row.

JHEP06(2026)274

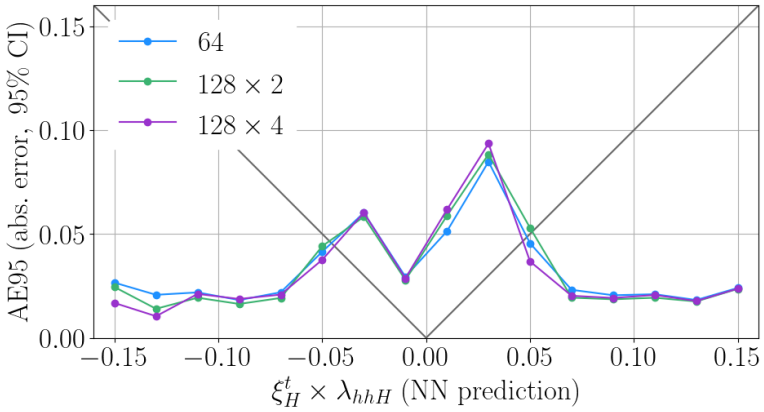

Figure 22. AE95 for different NN architectures. The one employed in this project is the simplest 1-hidden layer network with 64 neurons (blue), the more complex architectures include 2- and 4-hidden layer networks with 128 neurons each, shown in green and purple, respectively.

Data Availability Statement. This article has no associated data or the data will not be deposited.

Code Availability Statement. This article has no associated code or the code will not be deposited.

References

[1] ATLAS collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys. Lett. B* **716** (2012) 1 [arXiv:1207.7214] [INSPIRE].

[2] CMS collaboration, *Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC*, *Phys. Lett. B* **716** (2012) 30 [arXiv:1207.7235] [INSPIRE].

[3] ATLAS and CMS collaborations, *Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at* $\sqrt{s} = 7$ *and* 8 *TeV*, *JHEP* **08** (2016) 045 [arXiv:1606.02266] [INSPIRE].

[4] J. de Blas et al., *Higgs Boson Studies at Future Particle Colliders*, *JHEP* **01** (2020) 139 [arXiv:1905.03764] [INSPIRE].

[5] ATLAS collaboration, *A detailed map of Higgs boson interactions by the ATLAS experiment ten years after the discovery*, *Nature* **607** (2022) 52 [*Erratum ibid.* **612** (2022) E24] [arXiv:2207.00092] [INSPIRE].

[6] CMS collaboration, *A portrait of the Higgs boson by the CMS experiment ten years after the discovery.*, *Nature* **607** (2022) 60 [arXiv:2207.00043] [INSPIRE].

[7] H. Abouabid et al., *Benchmarking di-Higgs production in various extended Higgs sector models*, *JHEP* **09** (2022) 011 [arXiv:2112.12515] [INSPIRE].

JHEP06(2026)274

[8] J.M. Cline and P.-A. Lemieux, *Electroweak phase transition in two Higgs doublet models*, *Phys. Rev. D* **55** (1997) 3873 [hep-ph/9609240] [INSPIRE].

[9] L. Fromme, S.J. Huber and M. Seniuch, *Baryogenesis in the two-Higgs doublet model*, *JHEP* **11** (2006) 038 [hep-ph/0605242] [INSPIRE].

[10] J.M. Cline, K. Kainulainen and M. Trott, *Electroweak Baryogenesis in Two Higgs Doublet Models and B meson anomalies*, *JHEP* **11** (2011) 089 [arXiv:1107.3559] [INSPIRE].

[11] G.C. Dorsch, S.J. Huber, T. Konstandin and J.M. No, *A Second Higgs Doublet in the Early Universe: Baryogenesis and Gravitational Waves*, *JCAP* **05** (2017) 052 [arXiv:1611.05874] [INSPIRE].

[12] K. Kajantie, M. Laine, K. Rummukainen and M.E. Shaposhnikov, *Is there a hot electroweak phase transition at $m_H \gtrsim m_W$?*, *Phys. Rev. Lett.* **77** (1996) 2887 [hep-ph/9605288] [INSPIRE].

[13] G.C. Dorsch, S.J. Huber and J.M. No, *A strong electroweak phase transition in the 2HDM after LHC8*, *JHEP* **10** (2013) 029 [arXiv:1305.6610] [INSPIRE].

[14] G.C. Dorsch, S.J. Huber, K. Mimasu and J.M. No, *Echoes of the Electroweak Phase Transition: Discovering a second Higgs doublet through $A_0 \to ZH_0$*, *Phys. Rev. Lett.* **113** (2014) 211802 [arXiv:1405.5537] [INSPIRE].

[15] P. Basler et al., *Strong First Order Electroweak Phase Transition in the CP-Conserving 2HDM Revisited*, *JHEP* **02** (2017) 121 [arXiv:1612.04086] [INSPIRE].

[16] G.C. Dorsch, S.J. Huber, K. Mimasu and J.M. No, *The Higgs Vacuum Uplifted: Revisiting the Electroweak Phase Transition with a Second Higgs Doublet*, *JHEP* **12** (2017) 086 [arXiv:1705.09186] [INSPIRE].

[17] P. Basler, M. Mühlleitner and J. Wittbrodt, *The CP-Violating 2HDM in Light of a Strong First Order Electroweak Phase Transition and Implications for Higgs Pair Production*, *JHEP* **03** (2018) 061 [arXiv:1711.04097] [INSPIRE].

[18] D. Gonçalves, A. Kaladharan and Y. Wu, *Electroweak phase transition in the 2HDM: Collider and gravitational wave complementarity*, *Phys. Rev. D* **105** (2022) 095041 [arXiv:2108.05356] [INSPIRE].

[19] T. Biekötter et al., *The trap in the early Universe: impact on the interplay between gravitational waves and LHC physics in the 2HDM*, *JCAP* **03** (2023) 031 [arXiv:2208.14466] [INSPIRE].

[20] C. Caprini et al., *Science with the space-based interferometer eLISA. II: Gravitational waves from cosmological phase transitions*, *JCAP* **04** (2016) 001 [arXiv:1512.06239] [INSPIRE].

[21] C. Caprini et al., *Detecting gravitational waves from cosmological phase transitions with LISA: an update*, *JCAP* **03** (2020) 024 [arXiv:1910.13125] [INSPIRE].

[22] LISA Cosmology Working Group collaboration, *Gravitational waves from first-order phase transitions in LISA: reconstruction pipeline and physics interpretation*, *JCAP* **10** (2024) 020 [arXiv:2403.03723] [INSPIRE].

[23] T.D. Lee, *A Theory of Spontaneous T Violation*, *Phys. Rev. D* **8** (1973) 1226 [INSPIRE].

[24] J.F. Gunion, H.E. Haber, G.L. Kane and S. Dawson, *The Higgs Hunter's Guide*, *Front. Phys.* **80** (2000) 1 [INSPIRE].

[25] M. Aoki, S. Kanemura, K. Tsumura and K. Yagyu, *Models of Yukawa interaction in the two Higgs doublet model, and their collider phenomenology*, *Phys. Rev. D* **80** (2009) 015017 [arXiv:0902.4665] [INSPIRE].

JHEP06(2026)274

[26] G.C. Branco et al., *Theory and phenomenology of two-Higgs-doublet models*, *Phys. Rept.* **516** (2012) 1 [arXiv:1106.0034] [INSPIRE].

[27] T. Plehn et al., *Modern Machine Learning for LHC Physicists*, arXiv:2211.01421 [INSPIRE].

[28] M. Feickert and B. Nachman, *A Living Review of Machine Learning for Particle Physics*, arXiv:2102.02770 [INSPIRE].

[29] M. Abdughani et al., *Probing the triple Higgs boson coupling with machine learning at the LHC*, *Phys. Rev. D* **104** (2021) 056003 [arXiv:2005.11086] [INSPIRE].

[30] L. Alasfar et al., *Machine learning the trilinear and light-quark Yukawa couplings from Higgs pair kinematic shapes*, *JHEP* **11** (2022) 045 [arXiv:2207.04157] [INSPIRE].

[31] F. Arco, S. Heinemeyer, M. Mühlleitner and K. Radchenko, *Sensitivity to triple Higgs couplings via di-Higgs production in the 2HDM at the (HL-)LHC*, *Eur. Phys. J. C* **83** (2023) 1019 [arXiv:2212.11242] [INSPIRE].

[32] J.F. Gunion and H.E. Haber, *The CP conserving two Higgs doublet model: The Approach to the decoupling limit*, *Phys. Rev. D* **67** (2003) 075019 [hep-ph/0207010] [INSPIRE].

[33] S. Kanemura et al., *New physics effect on the Higgs selfcoupling*, *Phys. Lett. B* **558** (2003) 157 [hep-ph/0211308] [INSPIRE].

[34] S. Kanemura, Y. Okada, E. Senaha and C.-P. Yuan, *Higgs coupling constants as a probe of new physics*, *Phys. Rev. D* **70** (2004) 115002 [hep-ph/0408364] [INSPIRE].

[35] D.T. Nhung, M. Muhlleitner, J. Streicher and K. Walz, *Higher Order Corrections to the Trilinear Higgs Self-Couplings in the Real NMSSM*, *JHEP* **11** (2013) 181 [arXiv:1306.3926] [INSPIRE].

[36] C. Borschensky et al., *The trilinear Higgs self-couplings at $\mathcal{O}(\alpha_t^2)$ in the CP-violating NMSSM*, *Eur. Phys. J. C* **83** (2023) 118 [arXiv:2210.02104] [INSPIRE].

[37] S. Heinemeyer, M. Mühlleitner, K. Radchenko and G. Weiglein, *Higgs pair production in the 2HDM: impact of loop corrections to the trilinear Higgs couplings and interference effects on experimental limits*, *Eur. Phys. J. C* **85** (2025) 437 [arXiv:2403.14776] [INSPIRE].

[38] T. Plehn, M. Spira and P.M. Zerwas, *Pair production of neutral Higgs particles in gluon-gluon collisions*, *Nucl. Phys. B* **479** (1996) 46 [hep-ph/9603205] [INSPIRE].

[39] S. Dawson, S. Dittmaier and M. Spira, *Neutral Higgs boson pair production at hadron colliders: QCD corrections*, *Phys. Rev. D* **58** (1998) 115012 [hep-ph/9805244] [INSPIRE].

[40] A. Djouadi, W. Kilian, M. Muhlleitner and P.M. Zerwas, *Production of neutral Higgs boson pairs at LHC*, *Eur. Phys. J. C* **10** (1999) 45 [hep-ph/9904287] [INSPIRE].

[41] A. Arhrib et al., *Double Neutral Higgs production in the Two-Higgs doublet model at the LHC*, *JHEP* **08** (2009) 035 [arXiv:0906.0387] [INSPIRE].

[42] E. Asakawa et al., *Higgs boson pair production in new physics models at hadron, lepton, and photon colliders*, *Phys. Rev. D* **82** (2010) 115002 [arXiv:1009.4670] [INSPIRE].

[43] R. Grober, M. Muhlleitner and M. Spira, *Higgs Pair Production at NLO QCD for CP-violating Higgs Sectors*, *Nucl. Phys. B* **925** (2017) 1 [arXiv:1705.05314] [INSPIRE].

[44] P. Basler, S. Dawson, C. Englert and M. Mühlleitner, *Di-Higgs boson peaks and top valleys: Interference effects in Higgs sector extensions*, *Phys. Rev. D* **101** (2020) 015019 [arXiv:1909.09987] [INSPIRE].

[45] J. Baglio et al., *Full NLO QCD predictions for Higgs-pair production in the 2-Higgs-doublet model*, *Eur. Phys. J. C* **83** (2023) 826 [arXiv:2303.05409] [INSPIRE].

JHEP06(2026)274

[46] T. Biekötter et al., *First shot of the smoking gun: probing the electroweak phase transition in the 2HDM with novel searches for $A \to ZH$ in $\ell^+\ell^- t\bar{t}$ and $\nu\nu b\bar{b}$ final states*, *JHEP* **01** (2024) 107 [arXiv:2309.17431] [INSPIRE].

[47] H. Bahl et al., *HiggsTools: BSM scalar phenomenology with new versions of HiggsBounds and HiggsSignals*, *Comput. Phys. Commun.* **291** (2023) 108803 [arXiv:2210.09332] [INSPIRE].

[48] P. Bechtle et al., *HiggsBounds: Confronting Arbitrary Higgs Sectors with Exclusion Bounds from LEP and the Tevatron*, *Comput. Phys. Commun.* **181** (2010) 138 [arXiv:0811.4169] [INSPIRE].

[49] P. Bechtle et al., *HiggsBounds 2.0.0: Confronting Neutral and Charged Higgs Sector Predictions with Exclusion Bounds from LEP and the Tevatron*, *Comput. Phys. Commun.* **182** (2011) 2605 [arXiv:1102.1898] [INSPIRE].

[50] P. Bechtle et al., HiggsBounds – 4*: Improved Tests of Extended Higgs Sectors against Exclusion Bounds from LEP, the Tevatron and the LHC*, *Eur. Phys. J. C* **74** (2014) 2693 [arXiv:1311.0055] [INSPIRE].

[51] P. Bechtle et al., *Applying Exclusion Likelihoods from LHC Searches to Extended Higgs Sectors*, *Eur. Phys. J. C* **75** (2015) 421 [arXiv:1507.06706] [INSPIRE].

[52] P. Bechtle et al., *HiggsBounds-5: Testing Higgs Sectors in the LHC 13 TeV Era*, *Eur. Phys. J. C* **80** (2020) 1211 [arXiv:2006.06007] [INSPIRE].

[53] P. Bechtle et al., *HiggsSignals: Confronting arbitrary Higgs sectors with measurements at the Tevatron and the LHC*, *Eur. Phys. J. C* **74** (2014) 2711 [arXiv:1305.1933] [INSPIRE].

[54] P. Bechtle et al., *Probing the Standard Model with Higgs signal rates from the Tevatron, the LHC and a future ILC*, *JHEP* **11** (2014) 039 [arXiv:1403.1582] [INSPIRE].

[55] P. Bechtle et al., *HiggsSignals-2: Probing new physics with precision Higgs measurements in the LHC 13 TeV era*, *Eur. Phys. J. C* **81** (2021) 145 [arXiv:2012.09197] [INSPIRE].

[56] F. Arco, S. Heinemeyer and M.J. Herrero, *Sensitivity and constraints to the 2HDM soft-breaking Z_2 parameter m_{12}*, *Phys. Lett. B* **835** (2022) 137548 [arXiv:2207.13501] [INSPIRE].

[57] LHCb collaboration, *Studies of B-Hadron Decays to Charmless Final States at LHCb*, in the proceedings of the *Meeting of the APS Division of Particles and Fields*, Providence, U.S.A. (2011) [arXiv:1110.5535] [INSPIRE].

[58] M. Maniatis, A. von Manteuffel, O. Nachtmann and F. Nagel, *Stability and symmetry breaking in the general two-Higgs-doublet model*, *Eur. Phys. J. C* **48** (2006) 805 [hep-ph/0605184] [INSPIRE].

[59] V. Cacchio, D. Chowdhury, O. Eberhardt and C.W. Murphy, *Next-to-leading order unitarity fits in Two-Higgs-Doublet models with soft $\mathbb{Z}_2$ breaking*, *JHEP* **11** (2016) 026 [arXiv:1609.01290] [INSPIRE].

[60] ATLAS collaboration, *Combination of Searches for Resonant Higgs Boson Pair Production Using pp Collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector*, *Phys. Rev. Lett.* **132** (2024) 231801 [arXiv:2311.15956] [INSPIRE].

[61] M. Capozi and G. Heinrich, *Exploring anomalous couplings in Higgs boson pair production through shape analysis*, *JHEP* **03** (2020) 091 [arXiv:1908.08923] [INSPIRE].

[62] F. Arco, S. Heinemeyer and M.J. Herrero, *Sensitivity to triple Higgs couplings via di-Higgs production in the 2HDM at e^+e^- colliders*, *Eur. Phys. J. C* **81** (2021) 913 [arXiv:2106.11105] [INSPIRE].

[63] F. Arco et al., *Sensitivity to triple Higgs couplings via di-Higgs production in the RxSM at the (HL-)LHC and future e^+e^- colliders*, *JHEP* **06** (2025) 211 [arXiv:2502.03878] [INSPIRE].

JHEP06(2026)274

[64] F. Arco, S. Heinemeyer and M. Mühlleitner, *Large one-loop effects of BSM triple Higgs couplings on double Higgs production at e^+e^- colliders*, *JHEP* **01** (2026) 160 [arXiv:2505.02947] [INSPIRE].

[65] ATLAS collaboration, *Search for $H \to c\bar{c}$ and measurement of $H \to b\bar{b}$ in vector-boson fusion production with the ATLAS Detector*, arXiv:2511.21911 [INSPIRE].

[66] ATLAS collaboration, *Search for resonant pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state using pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Phys. Rev. D* **105** (2022) 092002 [arXiv:2202.07288] [INSPIRE].

[67] H. Bahl et al., *$\mathcal{CP}$-sensitive simplified template cross-sections for $t\bar{t}H$*, *JHEP* **10** (2024) 214 [arXiv:2406.03950] [INSPIRE].

[68] A. Paszke et al., *PyTorch: An Imperative Style, High-Performance Deep Learning Library*, arXiv:1912.01703 [INSPIRE].

[69] S. Gennai et al., *Search for heavy neutral MSSM Higgs bosons with CMS: Reach and Higgs-mass precision*, *Eur. Phys. J. C* **52** (2007) 383 [arXiv:0704.0619] [INSPIRE].

JHEP06(2026)274